TRANSCRIPTOMIC PROFILING OF HUMAN CELLS DESTINED FOR THERAPEUTIC APPLICATIONS

A Dissertation
Presented to
The Academic Faculty

by

Camila Medrano Trochez

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in Bioinformatics in the
School of Biological Sciences

Georgia Institute of Technology
May 2021

TRANSCRIPTOMIC PROFILING OF HUMAN CELLS DESTINED FOR THERAPEUTIC APPLICATIONS

Approved by:

Dr. Greg Gibson Advisor
School of Biological Sciences
Georgia Institute of Technology

Dr. Carolyn Yeago
Petit Institute of Bioengineering and Biosciences
Georgia Institute of Technology

Dr. King Jordan
School of Biological Sciences
Georgia Institute of Technology

Dr. Edwin Hortwitz
School of Medicine
Emory university

Dr. Jung Choi
School of Biological Sciences
Georgia Institute of Technology

Date Approved: January 22nd, 2021

To my family and friends

ACKNOWLEDGEMENTS

I want to start by thanking my PhD advisor Dr. Greg Gibson for his guidance and support. His passion about sciences kept me motivated during these years and inspire me to learn more about human genetics. I would also like to thank my committee members, Dr. King Jordan, Dr. Jung Choi, Dr. Ed Horwitz and Dr. Carolyn Yeago. Their support and feedback were critical for the success of my PhD research. It has been an incredible learning experience working in collaboration with Dr. Ed Horwitz and Dr. Carolyn Yeago.

To my friends and members of the Gibson's lab, thank you for the share knowledge about science and life. Specially to my friends Shashwat Deepali Nagar and Luz Medina Cordoba, thank you for your encouragement throughout my master and PhD at Georgia Tech.

I'm also deeply grateful to John Hiner for is unmeasurable patience and love, and his family that have supported me during this journey.

My Atlanta friends and the Georgia Tech swim club, thank you for the love and help throughout these years, and for helping me discover a different side of Atlanta and Georgia Tech, beside my lab.

Last, but not least, big thank you to my parents, Cristina Trochez and Edinson Medrano, and my brother and sisters, Ernesto, Claudia and Catalina Medrano Trochez. Your love and support allowed me to be here today. I would be forever grateful to have you as my family.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

DAMPs	danger-associated molecular patterns
PAMPs	pathogen-associated molecular patterns
NLM	National Library of Medicine
NK	Natural killer
NIH	National Institute of Health
GVHD	Graft-versus-host disease
HLA	Human leukocyte antigen
MHC	Major histocompatibility complex
APCs	Antigen-presenting cells
IL-1	Interleukin 1
IFNγ	Interferon γ
LPS	Lipopolysaccharide
Treg	Regulatory T cell
CTL	Cytotoxic T lymphocyte
ANAs	Antinuclear antibodies
OA	Osteoarthritis
CDC	Center of disease control
KL	Kellgren-Lawrence
MRI	Magnetic resonance imaging
TNFα	Tumor necrosis factor α
IL-6	Interleukin 6

IL-8	Interleukin 8
MSC	Mesenchymal stromal cells (also known as mesenchymal stem cells)
BMAC	Bone marrow aspirates concentrate
HSC	Hematopoietic stem cells
ISCT	Society for Cell and Gene Therapy
scRNA-seq	Single cell RNA sequencing

SUMMARY

Cell therapy is a growing field as many diseases are still untreatable. Different types of cells may be used as part of a treatment for a variety of diseases and conditions. As research advances, various cell types will be developed into treatments as novel cell therapies and studied for potential applications. Potential applications of cell therapies include treating cancers, autoimmune disease, urinary problems, and infectious disease, rebuilding damaged cartilage in joints, repairing spinal cord injuries, improving a weakened immune system, and helping patients with neurological disorders(1).

Numerous clinical trials have recently focused on the use of mesenchymal stromal cells (MSCs) as a cell therapy for various diseases with unmet medical challenges, including graft-vs-host disease, osteoarthritis, autism, and auto-immune diseases. MSCs are multipotent cells that have both regenerative and immunomodulatory capacity, and which are being developed for therapeutic intervention across a variety of inflammatory and immune conditions(2, 3). MSCs can be isolated from various sources, such as bone marrow, umbilical cord, adipose tissue and bone marrow aspirate concentrate (BMAC), which introduces tissue dependent variability between MSC-based cell products that also differ according to donor. Furthermore, manufacturing processes vary between sites (both clinical and commercial), leading to process-dependent variability. These sources of variability across the MSC field compound the ability to compare clinical trial results and have contributed to a lack of conclusive historical data to support the potential for clinical efficacy(4). Overall, there is an obvious need for deep phenotypic characterization of MSCs to compare heterogeneity as a function of tissue-of-origin as well as donor, and to

identify potential phenotypic signatures that can be used for biomarkers or quality attributes for MSC-based cell products.

My research focused on the study of MSCs as well as BMACs. BMACs contain small amounts of MSC but are still used in cell therapy as a MSC source. This study aims to assess the quality and gene expression profile of diverse cell preparation methods currently used in cell therapy, the transcriptional variability due to the tissue of origin of the cells used, and the pathways by which these cells modulate the immune system. I also evaluate the gene expression profiles of BMAC cell types from osteoarthritic (OA) patients and non-OA donors.

Recently, single-cell RNA sequencing (scRNA-seq) has emerged as one of the next generation cell characterization techniques that can be used to gain deeper insight into gene transcriptional signatures at the single-cell level(5, 6). scRNA-seq enables the examination of genomes or transcriptomes of individual cells, providing a high-resolution view of cell-to-cell variation or heterogeneity within a population. Moreover, this technique can be used to explore the distinct biology of individual cells and to understand temporal cellular processes and functions, such as differentiation, proliferation, and immune response potential(7).

Research advance 1: Chapter 2 characterizes the gene expression profiles of bone marrow and umbilical cord tissue-derived MSC. I assess the variability due to tissue of origin, batch effects, and donor effects. Samples derived from the same tissue separate in two major groups corresponding to high quality cells and low-quality cells. The high-quality cells further sub-divide in two subgroups referred to as high_a and high_b. Cells

from each donor belong to one or the other of these subgroups. The high-a cells are active cells expressing immunosuppressive genes, while high_b cells overexpress cell cycle genes, suggesting high_b cells may be more actively undergoing mitosis. The samples from the two different tissues also express qualitatively different gene expression profiles.

Research advance 2: Chapter 3 describes how MSC immunosuppressive capacities are enhanced at the transcriptional level by proinflammatory cytokines. I characterize the gene expression profiles of bone marrow derived MSC exposed to TNFα and IFNγ, showing that the MSC activated with IFNγ express higher levels of immunosuppressive genes. Additionally, RNAseq analysis performed on T cells exposed to IFNγ-enhanced MSC reveals enrichment in gene activity related to T cell suppression and cell stress related pathways.

Research advance 3: Osteoarthritis (OA) is a common degenerative disease, with no lasting cure. Even though it is considered a non-inflammatory disease, the immune system plays an important role in OA progression. The constant activation of pro-inflammatory cytokines due to tissue damage provokes deterioration of the cartilage. Cell therapy has been proposed as a treatment for this disease. MSC from different tissues of origin are being used to treat OA, such as commercial bone marrow derived MSC, bone marrow aspirates concentrate (BMAC) or autologous bone marrow derived MSC. Chapter 4 describes my characterization of the transcriptomic profiles of BMAC from OA patients and healthy donors. The results suggest overexpression of immune related pathways in BMAC samples from OA patients compared to samples from non-OA donors. Since pro-inflammatory cytokines play a role in progression of the disease, this study suggests the study of allogeneic BMAC from non-OA patients as possible option for OA therapy.

CHAPTER 1. INTRODUCTION

1.1. Immune system

The function of the immune system is to prevent or limit infection. The immune system can distinguish between healthy cells and unhealthy (damaged or dying) cells by recognizing danger-associated molecular patterns (DAMPs) present on the surface of unhealthy cells(8). Cells may be unhealthy because of infection or because of cellular damage caused by non-infectious agents like cancer. Infectious microbes such as viruses and bacteria release another set of signals recognized by the immune system called pathogen-associated molecular patterns (PAMPs)(8). These signals activate the immune system. Lack of activation results in the spread of the infection. On the other hand, when the immune system is activated when there is no need or is not turned off after the danger has passed, different problems arise, such as allergic reactions and autoimmune disease.

The immune system is complex, engaging numerous cell types that either circulate throughout the body or reside in a particular tissue. Each cell type plays a unique role, with different ways of recognizing problems, communicating with other cells, and performing specific functions. By understanding all the details behind this network, researchers may optimize immune responses to confront specific issues, ranging from infections to cancer.

The immune system can be broadly sorted into two arms: innate immunity and adaptive immunity. Innate immunity is the immune system we're born with. It consists of barriers and non-specific responses on and within the body that keep pathogens out. Components of the innate immune system include skin, mucus, stomach acid, enzymes

from tears and skin oils and the cough reflex(9). There are also chemical components of innate immunity, including substances called interferons and interleukins, and broad-specificity antimicrobial molecules(9). The adaptive, or acquired, immune system by contrast targets specific pathogens. Adaptive immunity is more complex than innate immunity. In this case, the pathogen must be processed and recognized by the body, and then the immune system creates antibodies specifically designed to target the pathogen(10). After the first exposure to the pathogen, the adaptive immune system remembers it, which allows for a more efficient response to the same pathogen in the future. In order to understand the immune system function and related diseases, it is critical to know the different components and their role within the immune system.

1.1.1. *Major components of the immune system*

All immune cells originate in the bone marrow and most of them develop into mature cells in the bone marrow. Some cell types, such as T cells, move out of the bone marrow to the thymus to complete their maturation process(11).

Skin: The skin is the first line of defence against microbes. Skin cells synthesize and secrete antimicrobial proteins.

Bone marrow: The bone marrow contains stems cells that can develop into a variety of cell types. Innate immune cells such as neutrophils, eosinophils, basophils, monocytes, dendritic cells, and macrophages come from myeloid stem cells present in the bone marrow. They are the first responders in case of infection. Mature myeloid and lymphoid immune cells are collectively called leukocytes. The adaptive immune cells, B cells and T cells, derive from lymphoid stem cells. These cells are responsible for specific responses

to microbes based on previous encounters. B cells make antibodies that attack bacteria and toxins, while T cells help destroy infected or cancerous cells. Killer T cells are a subgroup of T cells that kill cells that are infected or damaged. Helper T cells help determine which immune responses the body makes to a particular pathogen. Natural killer (NK) cells also derived from lymphoid progenitors and share features of both innate and adaptive immune cells: they provide immediate defences like innate cells, but also may be retained as memory cells like adaptive cells. B, T, and NK cells also are called lymphocytes.

Bloodstream: Immune cells constantly circulate throughout the bloodstream, patrolling for problems. When blood tests are used to monitor white blood cells, another term for immune cells, a snapshot of the immune system is taken. If a cell type is either scarce or overabundant in the bloodstream, this may indicate an immune problem.

Thymus: T cells mature in the thymus, a small organ located in the upper chest. In children, it supports negative selection against potentially self-reactive T-cells and thus prevents autoimmunity. The thymus shrinks in size in adults but is still used for positive selection of T-cells that recognize foreign cells.

Lymphatic system: The lymphatic system is a network of vessels and tissues composed of lymph, an extracellular fluid, and lymphoid organs, such as lymph nodes. The lymphatic system is a conduit for travel and communication between tissues and the bloodstream. Immune cells are carried through the lymphatic system and converge in lymph nodes, which are found throughout the body.

Lymph nodes: These are structures that produce and store cells that fight infections and disease. If the adaptive immune cells in the lymph node recognize pathogens brought

in from a distant area, they will activate, replicate, and leave the lymph node to circulate and address the pathogen. B-cells usually mature in the germinal centers of the lymph node.

Spleen: The spleen is an organ located behind the stomach and it is important for processing information coming from the bloodstream. Immune cells are enriched in specific areas of the spleen, and upon recognizing blood-borne pathogens, they will activate and respond accordingly. The spleen also helps control the amount of blood in the body and disposes of old or damaged oxygen-carrying red blood cells.

Mucosal tissue: Mucosal surfaces, like the respiratory tract and gut, are prime entry points for pathogens.

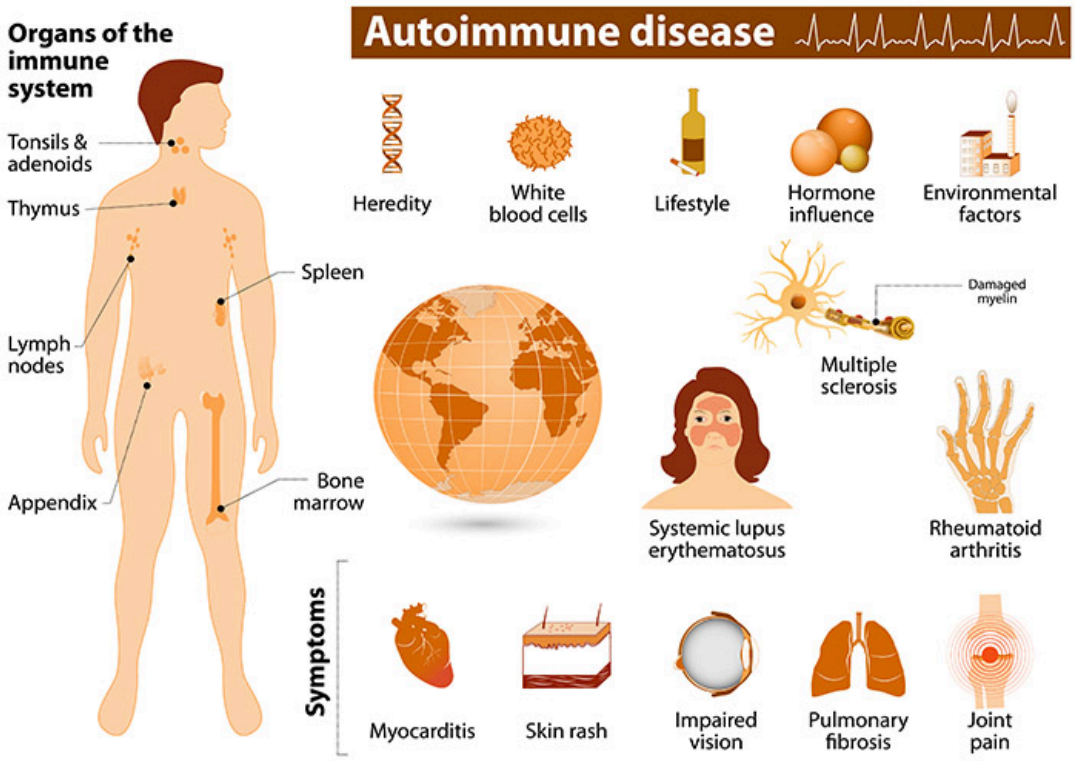


Figure 1. Immune system organs. Causes of autoimmune diseases and their symptoms. Figure published on the NIH website, 2020 (https://www.niehs.nih.gov/health/topics/conditions/autoimmune/index.cfm)

1.1.2. *Immune related diseases*

The immune system plays a major role in a daily fight against all pathogens. Multiple diseases can derive from unbalanced immune regulation. Disorders of the immune system can result in autoimmune diseases, inflammatory diseases and cancer, all reasons why its regulation is so important. Some of these diseases are due to the overreaction of the immune system, while other diseases occur for lack of response of the immune system, when facing a pathogen.

Autoimmunity results from a hyperactive immune system, mistaking normal tissues as if they were foreign. In common diseases such as asthma and allergies a harmless

material, like pollen, food or mould, is mistaken for a severe threat which activates the immune system. More severe autoimmune diseases include Hashimoto's thyroiditis, rheumatoid arthritis, diabetes mellitus type 1 and systemic lupus erythematosus.

According to the National Institute of Health (NIH), more than 24 million people in the United States suffer from autoimmune diseases(12). An additional eight million people have auto-antibodies, blood molecules that indicate a person's chance of developing autoimmune disease(13). These diseases are affecting an ever-increasing proportion of people for reasons unknown. Likewise, the causes of these diseases remain a mystery. Studies indicate these diseases likely result from interactions between genetic and environmental factors. Gender, race, and ethnicity characteristics are all linked to a likelihood of developing an autoimmune disease(13).

Immunodeficiencies, on the other hand, occur when the immune system is not strong enough, resulting in life-threatening infections. In humans, immunodeficiency can either be the result of a genetic disease such as severe combined immunodeficiency, or a process of immunosuppression, sometimes drug-induced(14).

Graft-versus-host disease (GVHD), is another immunological disorder that affects many organ systems, including the gastrointestinal tract, liver, skin, and lungs(15). GVHD is observed in 50% or more of patients receiving an organ transplant, including hematopoietic stem cell therapy. Since the graft contains T cells, there is a chance that the recipient expresses tissue antigens that are also present in the donor, as a consequence of which, the patient is incapable of mounting an effective response to eliminate the T

cells(16). Patients with a compromise immune system who receive white blood cells from another individual are at particularly high risk for this disease.

The donor T cells respond to human leucocyte antigens (HLAs) on recipient cells(17). HLAs, which are highly polymorphic proteins, are encoded by the major histocompatibility complex (MHC). Class I HLA (A, B, and C) proteins are expressed on almost all cells of the body. Class II proteins (such as DR, DQ, and DP) are mainly expressed on leukocytes. Their expression can also be induced on other cell types after inflammation or injury. The acute GVHD is related to the degree of mismatch between HLA proteins from donor and patient(18).

The disease results from a typical inflammatory process mediated by donor lymphocytes in the recipient. The immune system then attacks the cells displaying foreign HLA proteins, which in this case are self-cells. Levels of proinflammatory cytokines and chemokines increase, and the expression of key receptors on antigen-presenting cells (APCs) is stimulated, enhancing the presentation of the patient's HLA proteins to the donor immune cells that mediate GVHD(19, 20). Unchecked, GVHD can result in a sever and life-threatening autoimmune reaction to the organ transplant.

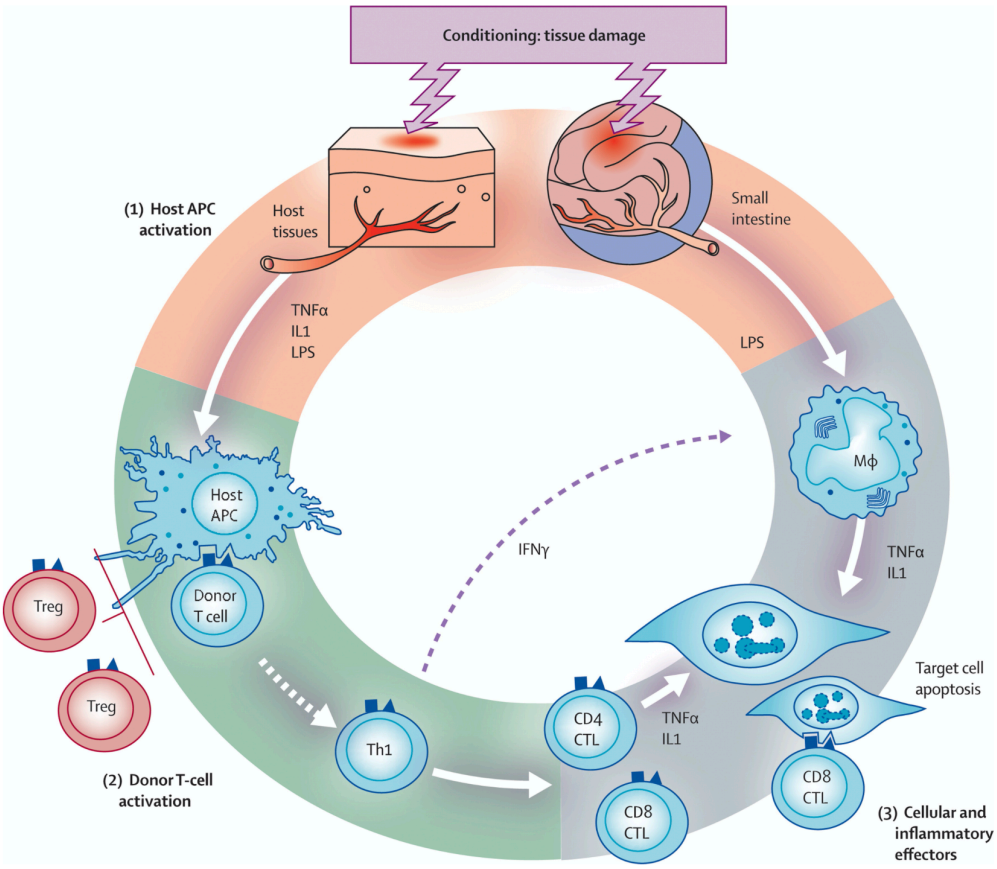


Figure 2. Pathophysiology of acute GVHD(15).

IL-1=interleukin 1. IFN γ=interferon γ. LPS=lipopolysaccharide. Treg=regulatory T cell. Th1=T-helper 1 cell. CTL=cytotoxic T lymphocyte.

1.1.3. *Diagnosis and treatment of immune related diseases*

More than 80 autoimmune diseases are known. Some of them are well known, such as type 1 diabetes, multiple sclerosis, lupus, and rheumatoid arthritis, while others are rare and difficult to diagnose. With rare autoimmune diseases, patients may suffer years before getting a diagnosis. Even though symptoms of immune diseases vary, fever and fatigue are common signs. Most of the time, immune deficiencies are diagnosed with blood tests that either measure the level of immune elements or their functional activity. Allergic conditions may be evaluated using either blood tests or skin allergy tests.

Most of these diseases have no cure. Some require lifelong treatment to ease symptoms. In autoimmune conditions, medications that reduce the immune response, such as corticosteroids, can be helpful.

In a study published in Arthritis and Rheumatology(21), the researchers found that the prevalence of antinuclear antibodies (ANAs), the most common biomarkers of autoimmunity, is significantly increasing in the U.S. overall and particularly in certain groups. These groups include males, non-Hispanic whites, adults 50 years and older, and adolescents. The study is the first to evaluate ANA changes over time in a representative sampling of the U.S. population. The reasons for this increase in ANA are not clear, but they are concerning and may suggest a possible increase in autoimmune disease cases in the future.

It is clear that we are far from completely understanding immune related diseases, which is why is so important not only to clarify the origin of these diseases but also to develop new strategies to treat or cure them.

1.2. Osteoarthritis

Osteoarthritis (OA) is a joint disorder involving cartilage, bone and synovial membrane. It can affect both small joints, such as toe and finger joints, as well as the large joints of the hip and knee(22). Globally, prevalence varies depending on the population genetics and environment, going from 17% to 47%(23, 24) . According to the Center for Disease Control and Prevention (CDC) OA affects over 32.5 million adults in the US(25). Its progress is usually slow; the pain and restricted function usually appears in people over 40 years old, and its prevalence rises with age.

Initially, osteoarthritis has been considered to be a disease of the articular cartilage, but recent research has indicated that the condition involves the entire joint(26-28). OA is classed as a degenerative disease due to the joint deterioration over time. The loss of articular cartilage is thought to be the primary change, but this is followed by a combination of cellular and biomechanical changes, including subchondral bone remodeling, formation of osteophytes (bone spurs), the development of bone marrow lesions, and meniscal damage, among other pathologies(29-32) (Figure 1.3). OA is considered to be a non-inflammatory condition even though inflammation is a symptom.

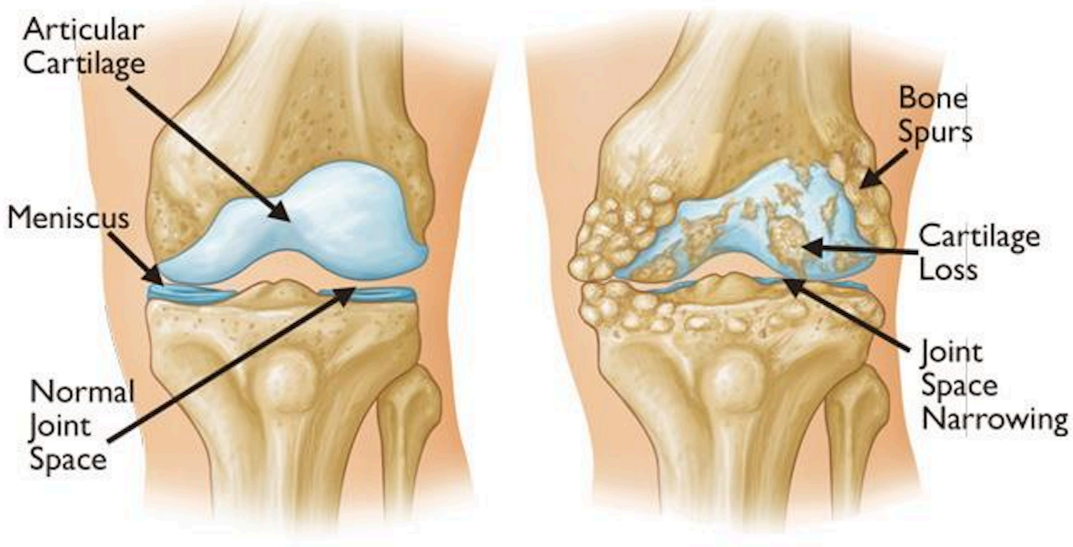


Figure 3. Joint anatomy in healthy and osteoarthritic patient, copied from https://orthoinfo.aaos.org/en/diseases--conditions/arthritis-of-the-knee, accessed Dec 20, 2020.

1.2.1. *Diagnosis of OA*

OA is diagnosed and graded by the degree of bone changes seen on an X-ray. The most commonly used X-ray tool to grade the severity of the disease is the Kellgren-Lawrence (KL) scale. This ranges from a score of zero, where no change is seen, to a

maximum score of four. A maximum score means the changes in the joint are severe, showing a significant decrease in joint space and marked osteophyte deposition.

1.2.2. *Pathophysiology and risk factors*

The following joint tissues are affected in OA.

Bone: In most people, bone changes in OA are seen much later in life, and the triggers are poorly understood. In some people, the composition of bone is affected by inherited disorders, which can lead to an onset of OA at an earlier age(22). The bone just beneath the cartilage in the synovial joint is known as subchondral bone and has a number of different zones. The upper zone, called the subchondral bone plate, is composed of relatively non-porous bone with a limited blood supply. Beneath this is a spongier layer of bone called cancellous or trabecular bone. In a healthy person, the bone is constantly modified by modelling and resorption. This allows it to repair and adapt to mechanical changes. In OA patients, the subchondral bone plate thickens, and osteophytes are formed at the joint margins. It is not yet clear whether changes within subchondral bone precede changes in the articular cartilage or whether they accompany disease progression. However, the two processes are closely related(33). MRI studies have demonstrated that these bone lesions themselves are associated with development and worsening of cartilage loss(34).

Synovial membrane: It's unclear whether the morphological changes that occur in the osteoarthritic synovial membrane are primary or whether they are the result of joint inflammation, cartilage degradation, and lesions of the subchondral bone(35). The synovial inflammation or synovitis is believed to be induced at first by the proteolytic degradation

components of the cartilage matrix. These components are released into the synovial fluid and are phagocytosed by macrophages. This process leads to the synthesis of mediators contributing to synovial inflammation and degradation of the cartilage(36). Studies of the changes in the synovium that occur at various stages of osteoarthritis have found that the amount of fibrin deposited in the synovial membrane and the degree of leukocyte infiltration are correlated with disease severity(37). High levels of osteopontin, a protein synthesize by various tissues like bone marrow cells and macrophages, has been correlated with disease severity(38).

Cartilage: Normal adult cartilage is formed by extracellular matrix (water, collagen, proteoglycans and a very small component of calcium salt) and chondrocytes(38). Pressure on cartilage causes the cartilage cells (chondrocytes) to produce a collagen and protein matrix. Replacement of the matrix is normally a slow process, but excessive loading, like obesity or genetic polymorphisms, can lead to activation of enzymes (metalloproteases) that digest it, leading to thinning and damage(39). Osteoarthritis result from failure of chondrocytes to maintain homeostasis between synthesis and degradation of these extracellular matrix components(40). It is not known what initiates the imbalance between degradation and repair of cartilage. The inflammation causes an increase in enzymatic activity that may allow the formation of "wear" particles, which could be engulfed by resident macrophages(33). When the production of these "wear" particles overwhelm the capacity of the immune system to eliminate them, they become mediators of inflammation, stimulating the chondrocyte to release degradative enzymes. This starts a cascade of actions that lead to the release of proinflammatory cytokines, like TNFα, IL-1 and IL-6. These cytokines will activate the chondrocytes which will lead to the secretion of

metalloproteinases and inhibition of type II collagen production, increasing cartilage degradation(41).

Inflammation: Once cartilage has been damaged, the products of that damage are engulfed and digested by synovial cells. This leads to the production of inflammatory mediators, such as pro-inflammatory cytokines interleukin IL-1, IL-6 and IL-8, tumour necrosis factor and prostaglandin PGE2.

The degree of inflammation is a key difference between OA and rheumatoid arthritis. OA is not always recognized as an inflammatory disease because there are often low levels of leukocytes in the synovial fluid, one of the parameters used to identify an inflammatory condition. Inflammation in OA is usually restricted to a specific area of a joint, but can sometimes be so pronounced that inflammatory chemicals are seen in the circulation as well as the synovial fluid. The degree of inflammation is associated with disease severity and pain.

Meniscus: Meniscal degeneration is commonly seen in osteoarthritis, where the meniscus appears torn, fragmented or completely destroyed(42). Degeneration of the meniscus initiates from within the tissue. Type I collagen content decreases gradually from the surface zone to the middle and the deep zone of osteoarthritic meniscus(43). Levels of proteoglycan also increase in osteoarthritic meniscus when compared to normal meniscus(43). All these intrameniscal changes correlate with peri-meniscal synovitis, and calcification (which is not limited to the outer, peripheral portion of the menisci(44, 45)) further contributes to meniscal degeneration and reduced meniscal strength. The meniscus

is then less able to stand loading and force transmission during normal movements of the joint, further leading to degenerative tears(46).

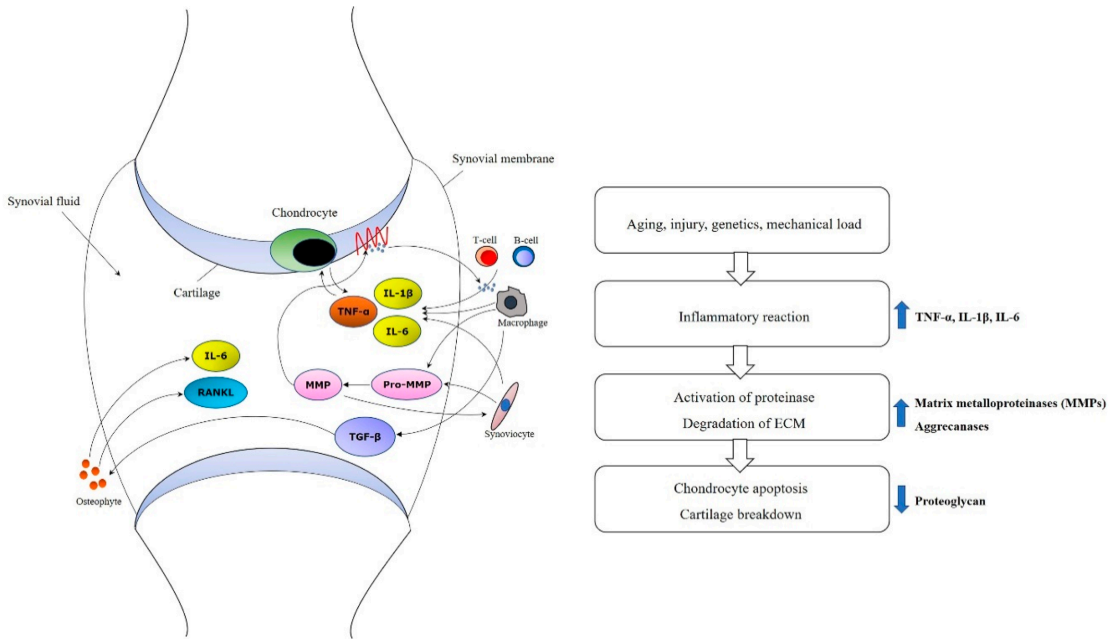


Figure 4. Molecular mechanism of osteoarthritis. Image published by Bhattacharyya et al. 2003(47).

1.2.3. *Treatment*

Current treatments target the pain (Nonsteroidal anti-inflammatory drugs and Acetaminophen) or the mobility (Physical therapy). There are no pharmaceutical treatments that can reverse the disease, so attention is turning to research on the effect of cell therapy for osteoarthritic patients. As of December 2020, there are 206 clinical trials for osteoarthritis involving cell therapy. Multiple cells types have been used, including mesenchymal stromal cells and bone marrow aspirates. Mesenchymal stromal cells (MSC) are widely used in clinical trials(48-50). Some clinical trials have seen encouraging

outcomes when using bone marrow aspirates concentrate, which also contains MSC(47, 51-54).

The reproducibility of the results seen in previous research is still challenging. A wide variety of cells are used for this purpose, introducing donor, technical, and tissue-fo-origin related variability. More studies need to be done to understand the mechanism by which these cells modulate the patient's condition.

1.3. Cell therapy

Cell therapy (also called cell transplantation) is a therapy in which viable human cells are injected, grafted or implanted into a patient in order to provide a medicinal effect. For example, in the case of ischemic heart disease, transplantation of therapeutic cells can improve perfusion, improve heart function and ultimately improve quality of life for patients(55). With new technologies many different types of cells may be used as part of a treatment for a variety of diseases and conditions. Some of the cells that may be used include hematopoietic stem cells (HSC), skeletal muscle stem cells, mesenchymal stromal cells, lymphocytes, dendritic cells, and cell aspirates (from bone marrow or stromal vascular fractions)(56).

Cell therapy trials have shown promising results for multiple disease treatments. Clinicaltrials.gov lists more than 35,000 studies for cell therapy assessment. Some cells, like MSC, have been shown to have little or no undesirable effects, suggesting that they are generally a safe option for therapeutic purposes(48-50, 57, 58). However, quality control is essential both to maintain safety and promote uniform potential. The cells injected to the patients may come from the patient him- or herself (autologous) or from an

unrelated healthy donor (allogeneic). Either way the rejection rate is of these cells is extremely low with appropriate medical care.

Numerous clinical trials with promising results have focused on the use of MSCs as a cell therapy for various diseases with unmet medical challenges, including graft-vs-host disease, osteoarthritis, autism, and auto-immune diseases(48-50, 57-60).

MSCs are multipotent cells that have both regenerative and immunomodulatory capacity, and which are being developed for therapeutic intervention across a variety of inflammatory and immune conditions(2, 3). MSCs can be isolated from various tissues, such as bone marrow, umbilical cord, and adipose tissue, which introduces tissue dependent variability between MSC-based cell products that also differ according to donor. Furthermore, manufacturing processes vary between sites (both clinical and commercial), leading to process-dependent variability. These sources of variability across the MSC field compound the ability to compare clinical trial results and have contributed to a lack of conclusive historical data to support the potential for clinical efficacy(4).

MSCs are typically identified based on the expression status of a panel of specific surface markers, their ability to adhere to plastic, and their tri-lineage differentiation potential to adipocytes, osteoblasts and chondroblasts, following recommendations of the International Society for Cell and Gene Therapy (ISCT)(61, 62). However, these minimal MSC identification and functional criteria, especially surface marker expression, often do not correlate with the regenerative or immunomodulatory functions of specific cultures(63). Moreover, the proportion of “stem” like cells that have high regenerative capability varies across MSC donors(64). Overall, there is an obvious need for deep

phenotypic characterization of MSCs to compare heterogeneity as a function of tissue-of-origin as well as donor, and to identify potential phenotypic signatures that can be used for biomarkers or quality attributes for MSC-based cell products.

As research advances, various cell types will be developed into treatments as novel cell therapies and studied for potential applications. Potential applications of cell therapies include treating cancers, autoimmune disease, urinary problems, infectious disease, rebuilding damaged cartilage in joints, repairing spinal cord injuries, improving a weakened immune system, and helping patients with neurological disorders(1). Even though, MSCs are widely used for cell therapy, other cell types, like hematopoietic stem cells (HSCs), or group of cells, like bone marrow aspirate concentrate and stromal vascular fraction, are also used for medical purposes.

1.4. Transcriptomics

The aim of this study is to assess the variability in function of MSC from different origins and BMAC from OA patients and non-OA donors. In order to do this analysis, I decided to use transcriptomics.

Transcriptomics is the study of the transcriptome, the complete set of RNA transcripts that are produced by the genome, in a specific cell, using high-throughput methods, such as microarray analysis, bulk RNA-seq and single cell RNA-seq.

RNA-Seq may be used to identify genes within a genome and quantify the activity of the genes at a point in time. The read counts can be used to accurately model the relative gene expression level of each gene, and the comparison of transcriptomes allows the

identification of genes that are differentially expressed in distinct cell populations, or in response to different treatments. RNA-Seq also provides a more precise measurement of levels of transcripts and their isoforms compare to other methods such as microarray analysis(65). In general, the transcriptome analysis is of growing importance in understanding how altered expression of genetic variants contributes to complex diseases such as cancer, diabetes, and heart disease. Analysis of genome-wide differential RNA expression provides researchers with greater insights into biological pathways and molecular mechanisms that regulate cell fate, development, and disease progression. However, like any other approach, transcriptomics has its limitations. mRNA abundance is an unreliable indicator of protein activity(66) as the transcript quantitation can be affected by biases created during cDNA library construction and sequence alignment.

Even though RNA-seq have limitations, it allows to understanding the transcriptome profile of different cells, which is essential for interpreting the functional elements of the genome and revealing the molecular compounds of cells and tissues(65).

CHAPTER 2. SINGLE CELL RNA-SEQ OF OUT OF THAW MESENCHYMAL STROMAL CELLS SHOWS STRIKING TISSUE OF ORIGIN DIFFERENCES AND INTER DONOR CELL CYCLE VARIATIONS

2.1. Abstract

Mesenchymal stromal cells (MSCs) from a variety of tissue sources are widely investigated in clinical trials, and the MSCs are often administered immediately after thawing the cryopreserved product. While previous reports have examined the transcriptome of freshly cultured MSCs from some tissues, little is known about the single-cell transcriptomic profiles of out-of-thaw MSCs from different tissue sources. Such understanding could help determine which tissue origins and delivery methods are best suited for specific indications. Here, we characterized cryopreserved MSCs, immediately post-thaw, from bone marrow (BM) and cord tissue (CT), using single-cell RNA sequencing (scRNA-seq). We show that out-of-thaw BM- vs. CT-MSCs have significant differences in gene expression. Gene-set enrichment analyses implied divergent functional potential. In addition, we show that MSC-batches can vary significantly in cell cycle status, suggesting different proliferative vs. immunomodulatory potentials. Our results provide a comprehensive single-cell transcriptomic landscape of clinically and industrially relevant MSC products.

2.2. Introduction

Mesenchymal Stromal Cells (MSCs), often referred to as Mesenchymal Stem Cells or Signaling Cells, are cells isolated from various tissues that have shown multipotent, regenerative, and immunomodulatory capacities in vitro. These cells, from a variety of tissue-sources, are being evaluated for therapeutic interventions, especially across a variety of inflammatory and immune conditions(2, 3). Numerous clinical trials have focused on the use of MSCs as a cell therapy for various diseases with unmet medical challenges, including graft-vs-host disease, osteoarthritis, autism, acute respiratory distress syndrome (ARDS), autoimmune diseases, and even COVID-19. A ClinicalTrial.gov search (date: August 7, 2020) with the keyword: MSC as other terms shows 4,044 studies that are either recruiting, not yet recruiting, enrolling by invitation, and active but not recruiting. MSCs are also widely used in developing engineered tissues ex vivo(67-69). Several are also working on developing MSC-based therapies and others are developing reagents and large-scale cell banks for eventual clinical use.

Despite such widespread interest in academia, clinical trials, and in industry, the characteristics of MSCs that are most correlative to their specific in vivo function remain unknown. MSCs can be isolated from various tissues, such as bone marrow, umbilical cord, placental, and adipose tissue, which introduces tissue-dependent variability between MSC-based cell products that may also differ according to donor. Furthermore, manufacturing processes vary between sites (both clinical and commercial), leading to process-dependent variability. These sources of variabilities across the MSC field confound the ability to compare clinical trial results and have contributed to a lack of conclusive historical data to support their potential for clinical use(4).

The International Society for Cell and Gene Therapy (ISCT) standards identify MSCs based on the expression status of a panel of specific surface markers, their ability to adhere to plastic, and their ex vivo tri-lineage differentiation potential to adipocytes, osteoblasts and chondroblasts(61, 62). However, these minimal MSC identification and functional criteria, especially surface marker expression, often do not correlate with their regenerative or immunomodulatory functions(63). Moreover, the proportion of "stromal" like progenitor cells that have high regenerative capability varies across MSC donors(64).

Therefore, there is an obvious need for deep phenotypic characterization of MSCs to compare heterogeneity as a function of tissue-of-origin as well as donor, and to identify potential phenotypic signatures that can be eventually used as predictive biomarkers or critical quality attributes (CQAs) for MSC-based products, and for understanding their putative Mechanisms of Action (MoAs).

Recently, single-cell RNA sequencing (scRNA-seq) has emerged as one of the next generation cell characterization techniques that can be used to gain deeper insight into gene transcriptional signatures at the single-cell level(5, 6). scRNA-seq enables the examination of genomes or transcriptomes of individual cells, providing a high-resolution view of cell-to-cell variation or heterogeneity within a population. Moreover, this technique can be used to explore the distinct biology of individual cells and to understand temporal cellular processes and functions, such as differentiation, proliferation, and immune response potential(7). scRNA-seq has been previously used to characterize hematopoietic differentiation(70-72) and immune cell subsets(73), including dendritic cells, monocytes(74), and innate lymphoid cells(75). A handful of reports have also used scRNA-

seq to characterize differential gene expression in freshly-prepared MSCs from umbilical cord(76), adipose tissue(62), Wharton's jelly(77) and bone marrow(78, 79).

In many clinical trial settings for allogeneic MSC-based off-the-shelf cellular therapies, to circumvent logistical and manufacturing challenges, MSC products are used post-thaw (directly from a frozen vial), rather than fresh (without freezing after culture), or culture-rescued (re-cultured after thawing)(3, 80). Since out-of-thaw MSC products could have different metabolic and functional characteristics from their fresh counterparts(80), phenotypic and functional characterization directly on the out-of-thaw MSC product is necessary to be able to find correlative attributes between their in vitro cell characteristics and corresponding clinical or pre-clinical efficacy. A comprehensive characterization of out-of-thaw MSC product from different tissue sources and donors at single-cell level may provide information on potential critical quality attributes (CQAs) and Mechanisms of Action (MoA) of the cells, and can be used to select MSC donor and/or sources for disease specific cell therapies.

In this study, we performed scRNA-seq using the drop-seq method(5, 81) to compare single-cell transcriptome profiles between commercially available bone marrow-derived MSCs (BM-MSCs) from six donors from RoosterBio Inc. (Frederick MD), and umbilical cord-tissue derived MSCs (UCT-MSCs) from four donors provided by Duke University. We characterized a total of 13 out-of-thaw samples from these ten MSC donors. Specifically, we assessed differences between individual donors as well as differences between MSC tissue sources. To overcome issues with zero counts that complicated differential expression analysis and to provide flexibility in normalization, we also

introduce a new analytical framework, scPool, in which similar cells from the same donor are pooled into pseudo-cells.

2.3. Results

2.3.1. *Donor Effects on Bone Marrow-Derived MSC Gene Expression*

A total of seven bone marrow-derived MSC (BM-MSC) samples from six donors were thawed and processed for scRNA-seq analysis (Tables 1 and 2). First, we compared two of the BM-MSC samples between pre-freeze and post-thaw conditions to understand freeze-thaw effects on MSC gene expression. Our analysis indicated a shift in the genetic profile between pre-freeze and post-thaw conditions (Figure S1). Pre-freeze samples showed significant overexpression of 1,743 genes relative to post-thaw samples at the 5% false discovery rate (FDR) threshold, while 310 genes were significantly overexpressed in the post-thaw samples compared to the pre-freeze samples. Some of the pathways overexpressed in the pre-freeze samples are cytokine signaling (FOS, MMP2, TLN1, FOSB), cell proliferation and cell adhesion (ZYX, ITGA5, CLIC1 etc.), while the pathways over-expressed in the frozen, post-thaw samples are carbohydrate interconversions (UGP2), cholesterol/Steroid biosynthesis and regulation of apoptosis (PSMA2, PSMB1). Having established that the freeze-thaw process imparts substantial changes in gene expression profiles of MSCs, we focused our analyses on post-thaw MSC products from BM and CT origins, since these are being widely used in numerous clinical trials.

Table 1 summarizes the sample data for all BM samples. Samples BM1 and BM2a were processed in laboratory A, while samples BM2b, BM3, BM4, BM5 and BM6 were processed in laboratory B (Table 1). An average of 305 cells were profiled per sample, with

an average read depth of 29,703, representing 12,348 UMI (Unique Molecular Identifier) and 3,836 expressed genes per cell (Table 1). The profiles were clustered with both Seurat(82) and SC3 pipelines(83). Since the latter is optimized for relatively small experimental designs, we present the results of SC3, but note that similar findings were obtained with Seurat (Figure S2). Before characterizing differential expression among samples, we confirmed that the MSC-identity markers established by the ISCT, namely NT5E (CD73), THY1 (CD90) and ENG (CD105) were detected in the majority of cells in the bone marrow and umbilical cord tissue derived MSCs (Figure S3). Furthermore, CD34, CD14, CD19 and PECAM1 - all markers of hematopoietic or lymphoid lineages, were absent.

Table 1. Number of cells, average number of genes and average number of UMI per cell per BM-MSC sample.

Samples	Lab	Number of Cells	Number of Reads	Mean Reads per cell	Median genes per cell	Number of cells post filtering	Sex	Age	Average nGenes per cell	Average nUMI per cell
BM1	A	293	29,671,530	15,469	2,218	293	Male	25	2,693	7,581
BM2a	A	310	29,943,109	46,467	3,855	252	Male	21	3,714	12,242
BM2b	B	199	41,511,093	39,417	4,686	199	Male	21	4,395	15,506
BM3	B	590	35,842,779	26,580	4,272	452	Male	22	4,193	13,415
BM4	B	557	25,262,984	17,668	3,803	542	Female	26	3,430	9,718
BM5	B	149	18,160,926	32,928	4,201	122	Male	31-45	4,057	13,591
BM6	B	317	21,841,113	29,393	4,554	275	Male	18-30	4,367	14,385

Projecting each cell against the first two Principal Components (PC) of gene expression, three clusters of single cell profiles were observed (Figure 5A). PC1 is highly negatively correlated with the total UMI count per cell. Accordingly, the smallest cluster located to the right, consists of 17% of the cells all of which had low UMI counts, typically fewer than 1,000 detected genes, and low ribosomal protein transcript counts (Figure 5B). These low UMI counts cells were more prevalent in two donors studied - one from each of the two laboratories (Lab A and Lab B; BM1 and BM4, respectively: Figure 5C), suggesting the low UMI count may not be related to the lab in which they were manufactured. It is not clear whether the unusual profile of these cells is a technical artefact, or has a biological basis, but they appear to be of low quality and were excluded from all subsequent analyses.

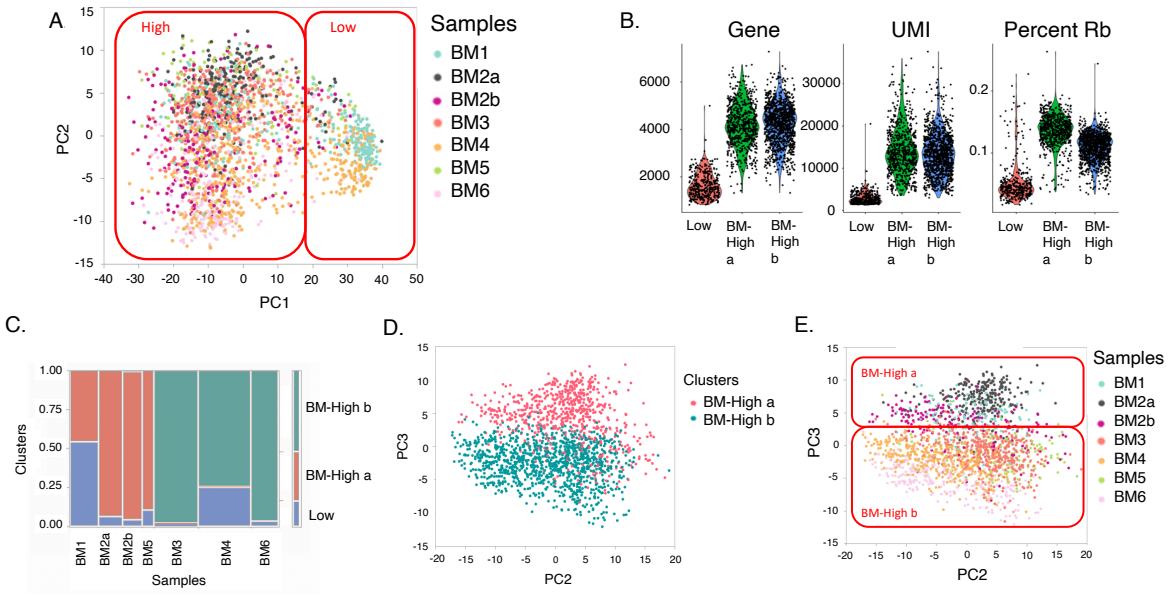


Figure 5. Clusters of BM-MSC transcriptome profile.

(A) The first two Principal Components of gene expression identify two broad clusters of cells, which are colored by sample: Cluster BM-Low, which correspond low UMI count cells and cluster BM-High, with high UMI count cells. (B) Violin plots show the density of the number of Genes, UMI, and Ribosomal Protein transcripts (RP) per cell. (C) Association of cells with clusters. The width of each column is proportional to the number

of cells in the indicated sample, and the color of each box corresponds to cells in cluster BM-Low (blue), BM-High a (red) or BM-High b (green). (D) Within cluster High, SC3 identifies two clusters of cells, which separate along PC3 as indicated by the red and blue points. (E) Shading of cells by sample confirms that cells from each donor belong to one of the two sub-clusters, although with subtle separation associated with PC3.

Two clusters identified by SC3 in the remaining high-quality datasets largely differentiate along PC3 (Figure 5D). Three of the five samples expanded in laboratory B (BM3, BM4 and BM6) were predominantly found in cluster BM-High_b; the other two samples along with both of the samples expanded in laboratory A (BM1, BM2a, BM2b and BM5) were predominantly found in cluster BM-High_a (Figure 5E). Both samples from the donor whose cells were cultured in each of the laboratories are in cluster BM-High a (BM2a and BM2b), suggesting that the difference is more likely to be donor-related than due to a laboratory or technical effect. Nevertheless, Figure 5E shows that even between the two laboratories, the cells from this donor tend to separate along PC3.

2.3.2. *Single Cell Pooling Enhances Differential Expression Analysis for Bone Marrow-MSC samples*

Although the above analysis follows current standard practice, it is nevertheless subject to the caveat that for a large percentage of the genes, transcripts are only observed in fewer than half of the cells, and consequently the differential expression analysis is mostly based on presence-versus-absence. In order to lend robustness to the conclusions, we elected to analyze differential expression with an alternative strategy that accounts for the impact of a high proportion of drop-outs in scRNA-seq. Our method, "scPool", is based on pooling of cells of the same type within samples, which ensures that a high proportion of genes are represented by an approximately normal count distribution. It also allows for

fitting of normalization strategies initially developed for bulk microarray or RNA-seq analysis.

After normalizing gene expression values to counts per 10,000 UMI, differential expression analysis between clusters BM-High_a and BM- High_b was performed in EdgeR with donor as a random covariate, yielding 4,624 genes at a FDR of 5%. There were 2,230 genes upregulated in cluster 2a, and 2,394 upregulated in cluster 2b (Figure 6A). Gene ontology analysis detected strong enrichment for multiple pathways involved in cell cycle regulation in cluster BM-High_b (Figure 6B). By contrast, cluster BM-High_a showed upregulation of multiple pathways related to immune signaling and other processes expected to be characteristic of functional MSCs (Figure 6B). On the basis of the cell cycle gene expression, cells in cluster BM-High_b may be preparing for or undergoing cell cycle division, whereas the cluster BM-High_a MSCs may be more likely to be in G0 phase. Alternately the two populations may simply express cell cycle related genes at different levels, without this reflecting cell cycle stages.

To implement the scPool strategy, we created random pools of pseudo-cells consisting of the sum of raw read counts for random draws of 20 cells (excluding the low-quality cluster 1 cells) within a donor's sample. The number of such pseudo cells ranged from 6 for sample BM5 to 20 for sample BM4. We then retained all genes with at least one read in 90% of the pseudo cells, a total of 6,422 genes. This dataset was normalized using the supervised normalization of microarrays (SNM) protocol(83) with cluster sub-type as the biological variable, adjusting for donor effects, and analysis of variance was used to detect differentially expressed genes. The procedure was repeated ten times, and the fold change and p-values were averaged to generate a robust list of cluster-specific

genes. Compared with the single cell analysis, 1,290 more significant differentially expressed genes were detected. The single cell analysis was done using the Wilcoxon rank sum test for the differentially expressed genes and a threshold lower than 0.05 for the adjusted Pvalue. Figure 6C shows generally higher significance than the single cell-level analysis, without over-estimation of fold-changes for a large fraction of the less-significant genes.

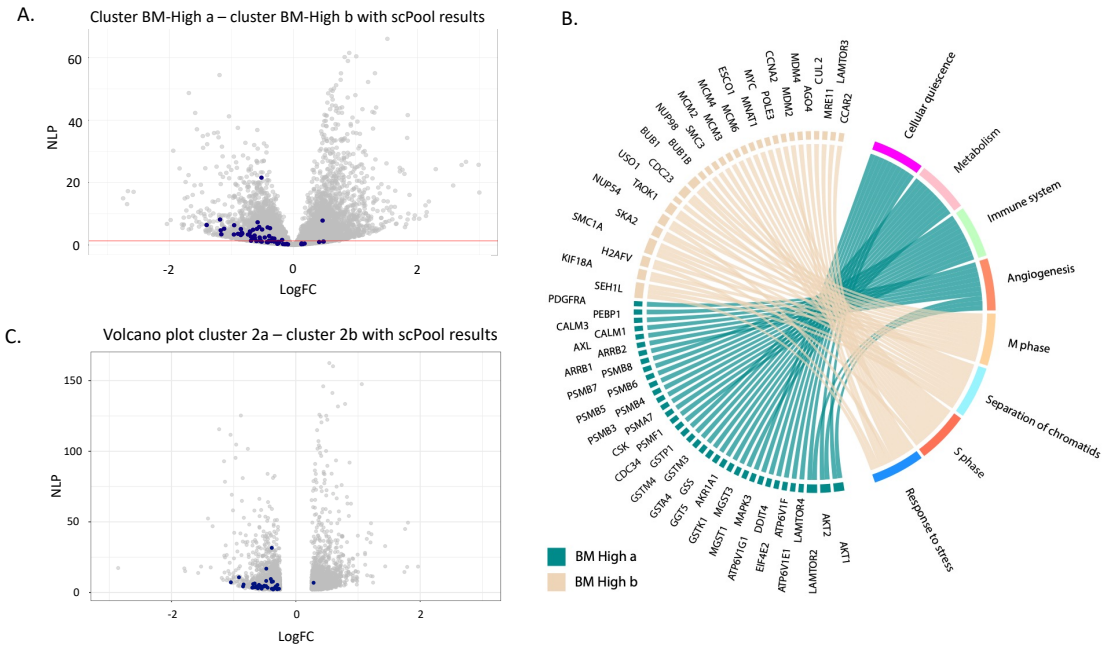


Figure 6. Differential expression between the two High clusters of Bone Marrow-MSC samples.

(A) Volcano plot of negative log P-value (NLP) against fold change, created with standard single cell level computation in edgeR. (B) Corresponding volcano plot based on the pseudo pools results. Blue shading indicates cell cycle markers obtained from cycleBase(84). (C) Chord diagram of gene ontology analysis highlighting the top 10 differentially expressed genes in each of 8 pathways representative of the up- and down-regulated genes. Ribbons link genes on the left to pathways on the right; genes associated with multiple pathways bifurcate. Note that in this depiction, the direction of differential expression is the same for all genes in the pathway.

Next, we probed the magnitude of donor contributions to variance within clusters, by performing analysis of variance with donor as the fixed effect of interest. Within just

the high-quality cluster BM-High_a cells, donor effects accounted for 8.5% of the variance. A similar result was observed for cluster BM-High_b.

2.3.3. *Donor Effects on Umbilical Cord Tissue Derived MSC Gene Expression*

Umbilical cord tissue derived MSC (UCT-MSC) samples from four donors were used for scRNA-seq analyses. A total of six scRNA-seq samples were prepared, all from the same lab, including three biological replicates of one donor sample (UCT1a, UCT1b and UCT1c). An average of 349 cells were profiled per sample, with an average read depth of 27,417, representing 9,700 UMI and 3,057 expressed genes per cell (Table 2). The UCT-MSC gene expression profiles were also analyzed with the SC3 pipeline.

Table 2. Number of cells, average number of genes and average number of UMI per cell per UCT-MSC sample.

Samples	Lab	Number of Cells	Number of Reads	Mean Reads per cell	Median genes per cell	Number of cells post filtering	Sex	Age	Average nGenes per cell	Average nUMI per cell
UCT1a	C	161	16,255,153	19,563	2,472	161	Male	--	2,808	8,458
UCT1b	C	251	31,333,645	38,809	3,840	251	Male	--	3,667	12,075
UCT1c	C	251	31,333,645	38,809	3,840	251	Male	--	2,477	7,321
UCT2	C	692	59,783,150	25,248	2,979	547	Male	--	3,053	9,381
UCT3	C	397	43,757,708	26,241	2,724	396	Male	--	2,720	9,262
UCT4	C	284	20,167,531	29,162	3,717	259	Male	--	3,616	11,711

Two major clusters of single cell profiles were again observed in the projection of the first two principal components of the UCT-MSC data (Figure 7A).

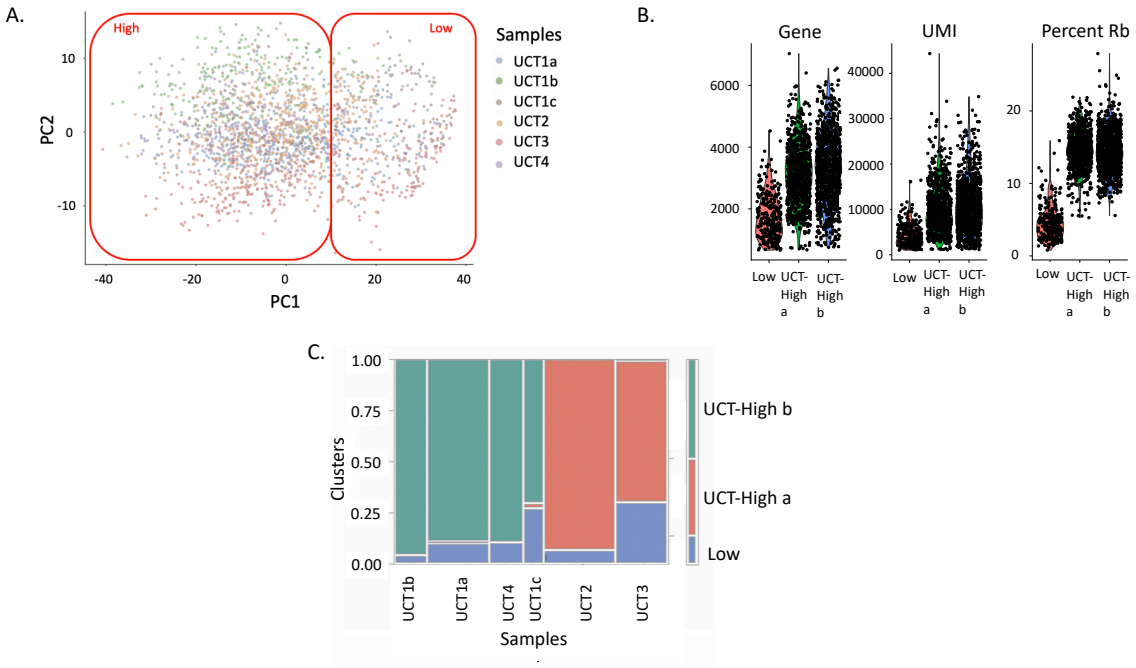


Figure 7. Clusters of UCT-MSC Profiles.

(A) The first two Principal Components of gene expression identify the two major clusters of cells, which are colored by sample. Most cells of each sample cluster together. (B) Violin plots show the density of the number of Genes, UMI, and Ribosomal Protein transcripts (RP) per cell. (C) Association of cells with clusters. The width of each column is proportional to the number of cells in the indicated sample, and the color of each box corresponds to cells in cluster 1 (blue), 2a (red) or 2b (green).

The smallest of these, consisting of 13% of the cells, was characterized by cells with low UMI counts, typically fewer than 2,000 detected genes (Figure 7B), similar to the BM-MCS analysis. These low UMI-count cells were present in every sample but again were more prevalent in two of the samples (UCT1c and UCT3: Figure 7C). It is not clear whether the origin of these cells is a technical artefact, or has a biological basis, but they also appear to be of low quality and were again excluded from all subsequent analyses. Within the high-quality cells, SC3 once more identified two clusters of cells, UCT-High_a

(UCT2 and UCT3) and UCT-High_b (UCT1a, UCT1b, UCT1c and UCT4), though in this case they did not clearly correspond to one of the Principal Components. The three replicates of donor UCT1 were primarily captured within the UCT-High_b cluster, suggesting consistency of technique.

Implementation of the scPool strategy for detecting differential gene expression between the two UCT-MSC clusters, after removing the low-quality cells, detected 2,526 genes at an FDR of 5%. Directional up-regulation of established marker genes for mitosis is evident in cluster UCT-High_b as indicated by blue points in the volcano plot Figure 8A. Gene ontology analysis (Figure 8B) indicates enrichment for up-regulation of collagen biosynthesis, integrin signaling, extracellular matrix (ECM) organization, and protein translation pathways in the cluster UCT-High_a MSCs, whereas the cluster UCT-High_b cells are enriched for cell cycle regulation, degradation of mitotic proteins, as well as various processes related to cell cycle progression, including CDC20 mediated degradation of Securin, and auto degradation of CDH1, suggesting potential donor-dependent heterogeneity in the gene expression profile of UCT-MSC.

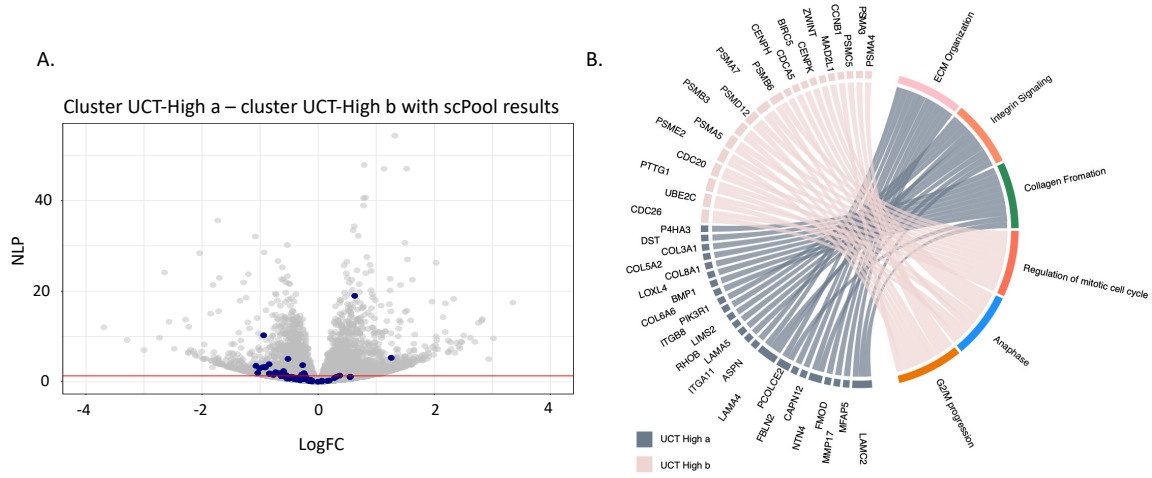


Figure 8. Differential expression between two high quality UCT-MSC clusters.

(A) Volcano plot of significance against fold difference in gene expression for the comparison of clusters 2a and 2b. Blue points indicate genes with established roles in cell-cycle regulation. (B) Chord diagram of gene ontology analysis highlighting the top 10 differentially expressed genes in each of 10 pathways representative of the up- and down-regulated genes. Ribbons link genes on the left to pathways on the right; genes associated with multiple pathways bifurcate.

2.3.4. *Comparison between Bone-Marrow and Cord-Tissue derived MSC single-cell gene expression profiles*

Direct comparison of MSCs derived from bone marrow and MSCs from umbilical cord tissue was performed by combining the analyses of the previous two datasets. As expected, Principal Component Analysis (PCA) of the raw single cell profiles again identified two major clusters of low and high UMI-abundance cells along PC1, but in this case PC2 of the joint analysis cleanly differentiates the BM and CT samples (Figure 9A). This result implies that there are significant differences in gene profile between MSCs derived from the two source tissues.

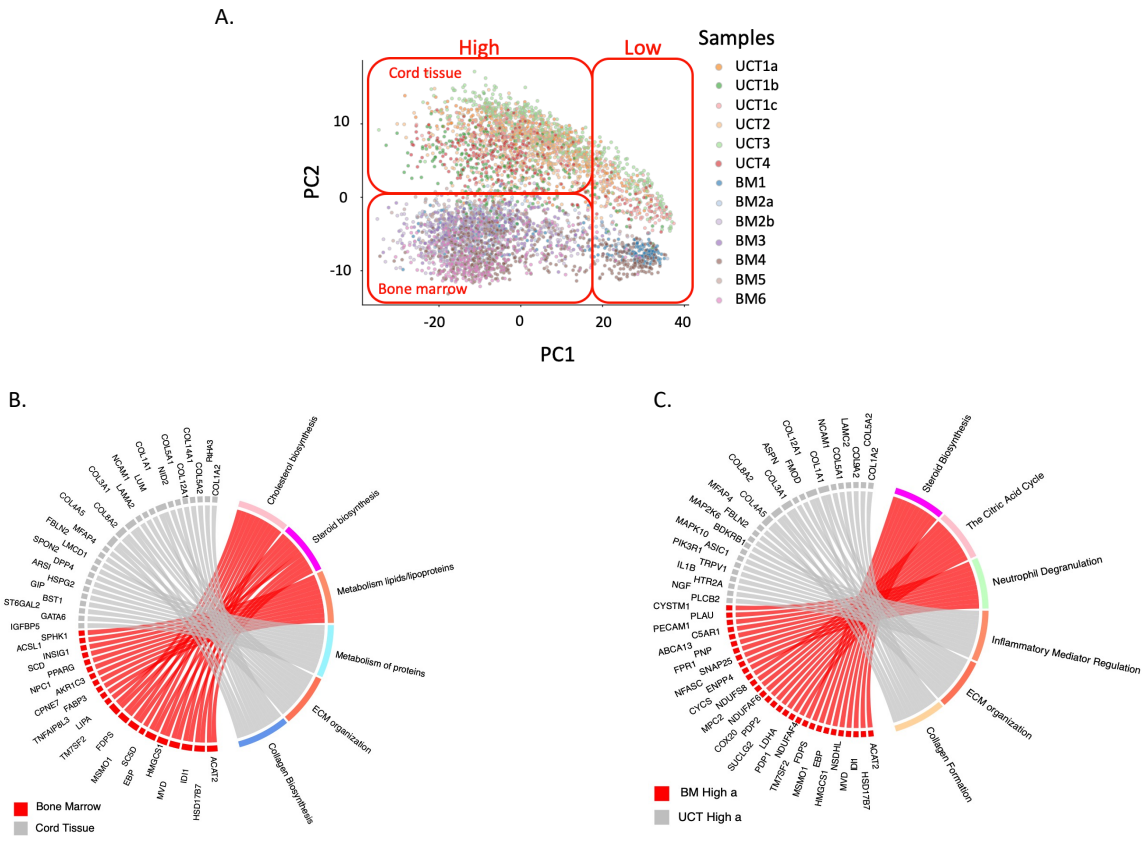


Figure 9. Differential expression analysis between high quality BM-MSC and UCT-MSC.

(A) Principal component analysis showing clustering of MSC samples by tissue of origin (bone marrow vs umbilical cord tissue). (B) Chord diagram summarizing differential expression between UCT-MSC and BM-MSC. (C) Cluster-specific pathway expression is not conserved between the two High clusters from the two tissues of MSC origin. Chord diagram contrasting clusters BM-High a versus UCT-High a.

We implemented scPool to identify differentially expressed genes between MSCs derived from the two sources, after removing the low-quality cells identified in our first series of analysis. Pools of 20 cells were again assembled computationally within donor and mixed model analysis of variance was performed, with donor as a random effect. 3,437 genes were found to be up-regulated in the BM-MSC, and 3,250 genes down-regulated, compared to UCT-MSCs. We highlighted the top 16 differentially expressed genes (Figure

S4). Gene ontology analysis of all the significantly differential expressed genes indicates enrichment for metabolism of lipids and lipoproteins, cholesterol biosynthesis, mitochondrial translation, and metabolic pathways in the BM-MSCs, whereas the UCT-MSCs were enriched for ECM organization, collagen biosynthesis and signal transduction (Figure 9B). UCT-MSC also showed relative up-regulation of mitotic cell cycle pathways, but this likely reflects the greater ratio of UCT-High_b to UCT-High_a cluster cells than of BM-High_b to BM-High_a cells, rather than a consistent trend favoring cell division in the UCT-MSC.

Pathway enrichment analysis also showed that the differences between the two clusters in each tissue are not consistently maintained. The chord diagram (Figure 9C) highlights pathways overexpressed in BM-High_a and UCT-High_a, which are not the same. These data confirm differences in the gene expression of non-dividing cells as a function of tissue of origin and suggest that the two types of MSCs are likely to have divergent regulatory and functional potentials. Importantly, the UCT-High_a population exhibit higher expression of genes involving pro inflammatory mediation, ECM organization and collagen biosynthesis, whereas the BM-High_a population had higher expression of steroid biosynthesis, the citric acid cycle and neutrophil degranulation genes (Figure 9C).

Next, we examined the expression level of genes that play important roles in the immunomodulatory response induced by MSCs. Focused comparison of expression of genes that are associated with cell adhesion, migration, immunosuppression and immunostimulation between the BM- and UCT-derived MSCs suggests tissue-of-origin and donor differences in gene activity (Figure S5). BM-High_a cells characteristically

overexpress transcripts encoding the membrane proteins prostaglandin synthase (PTGES2) and Endoglin (ENG) as well as the lysosomal protein CD63, compare to BM-High_b, whereas BM-High_b cells overexpress the genes CD46 (encoding a complement cofactor), CD47 (an integrin-associated protein), and CD146 (MCAM, cell adhesion molecule) , compare to BM-High_a. These genes are in general overexpressed in BM derived MSC compare to UCT derived MSC. BM derived MSCs have higher expression of the cell surface glycoprotein coding genes CD44 and CD59, as well as the nucleotidase NT5E and immune checkpoint molecule CD276 compare to UCT derived MSC. Conversely, when comparing UCT-High_a and UCT-High_b clusters, UCT-High_a cells overexpress the tetraspanin regulators of motility CD151, and cell surface protein coding genes CD99, THY1 and CD248, while UCT-High_b overexpressed CD9. The cell surface protein coding gene CD81 is not significantly differentially expressed between the clusters UCT-High_a and UCT-High_b. We looked at two genes associated with immunostimlation: CCL2 and CD109, which are overexpressed in UCT and BM derived MSCs, respectively.

We also looked at multiple pluripotent and stemness genes (Figure S6), none of which were found to be significantly differentially expressed between bone marrow and umbilical cord tissue.

2.4. Discussion

MSCs from various tissue sources are the subject of 4044 registered clinical trials (ClinicalTrials.gov- keyword: MSC as other terms with filters - not yet recruiting, recruiting, enrolling by invitations, and active, not recruiting; search date – August 7, 2020). It is thus important to develop robust high-throughput approaches for

characterization of diverse batches from various tissue sources in order to help evaluate reasons for success or failure of individual trials or patient responses. Single cell RNA sequencing is a relatively unbiased approach to profile the molecular attributes of individual cells. Potential utility of scRNA-seq includes characterization of heterogeneity that cannot be observed with bulk RNA-seq, and monitoring of the effect of the stage of the cell-cycle on transcriptional diversity.

Contrasting pre-freeze and post-thaw samples from 2 donors, we identified numerous differentially expressed genes that are associated with different types of cellular functions, such as cytokine signaling, cell proliferation, cell adhesion, cholesterol/steroid biosynthesis, and regulation of apoptosis. Previuosly, functional differences between pre-freeze (fresh) and post-thaw MSCs were also reported by others(85). In this study, however, we focused on in-depth scRNA-seq analysis of post-thaw MSCs as they are currently being tested as cell therapy products in many clinical trials.

Here we describe a droplet-based scRNA-seq comparison of donor, tissue-of-origin, and expansion conditions of out-of-thaw MSC variability, concluding that bone marrow and umbilical cord tissue-derived MSCs have significant differential expression that likely explains some of the documented differences between them, and that donor differences are modest yet significant. To our knowledge, six other scRNA-seq studies(5, 6, 76-79) of MSCs have been published, and our results are broadly concordant though with some important differences in emphasis. Barrett et. al., 2019(78) used a version of SmartSeq to deeply profile 103 Wharton's jelly-derived umbilical cord MSCs and 63 bone marrow-derived MSCs, identifying 463 differentially expressed genes enriched for activity in numerous processes including the matrisome, coagulation, angiogenesis, and wound-

healing via immune-regulation. The current study similarly finds a difference between cord tissue and bone marrow-derived MSCs. Additionally, our data also shows a cell cycle variability which seems to be related to the donor(78) (22).

Each of the other studies (5, 6, 76-79) has noted that the cell cycle gene expression is a major source of heterogeneity within donors. According to the Huang et. al., 2019(76) study, it appears the cell cycle is related to the immune regulatory potency of the MSCs. Previous work(77) has used a core set of G1/G2M/S markers to assign cells to each phase, and regressed out this source of variance before performing downstream analysis. We eschewed this approach both because of concerns over the reliability of the assignments, and to emphasize that the proportion of cells with low expression of these genes is an important component of among-donor differences in both BM-MSC and UCT-MSC. Reported higher proliferative capacity of Wharton's jelly-derived MSCs(78) is consistent with the higher proportion of mitotic genes in our UCT-MSCs relative to BM-MSCs. However, it should be emphasized that higher overall expression may not correlate with higher rates of proliferation, since expression levels may vary among donors without implying that a different fraction of cells are undergoing division. On the other hand, it appears that putative G0 cells that do not express cell cycle genes have quite different transcriptional properties that are directly relevant to their biological functions such as immunomodulatory potential. We note that each of our samples was profiled at population doubling level (PDL) ranged from 12-15, eliminating passage number as a source of variability in our study.

Other authors have also chosen to regress out "batch" effects before searching for heterogeneity, even though in each case "batch" appears to be coincident with "donor" (62,

77) or “Passage” (76). In the absence of biological replication, that is, two MSC preparations obtained independently from the same donor, it is impossible to know whether differences between sample populations have a biological or technical basis. Nevertheless, we estimate from principal component variance analysis that less than 10% of the overall expression variability is among donors/batches within each of the two clusters observed in both the BM- and UCT-MSC datasets (Figure 5C, Figure 7C). We see this minor source of variability is donor-related in the two instances where we had technical replicates from the same donor (in the case of the two BM-MSC samples cultured in different laboratories). The cells strongly tended to be assigned to the same sub-cluster BM-High_a or UCT-High_a. Whether or not these differences impact MSC function in clinical applications remains to be seen, additional large-scale comparisons with a large set of samples with high quality data on patient outcomes will need to be analyzed.

In this study the transcriptomes of human bone marrow and cord tissue-derived MSCs were analyzed via drop-seq single cell RNA-seq. Using this approach, new information about MSCs emerges. First, the differences between bone marrow-derived MSCs and cord-tissue derived MSCs were seen. Surprisingly, pathways up-regulated in G0 bone marrow-derived MSCs did not correspond to the same pathways upregulated in G0 cord tissue -derived MSC (Figure 9C). Further, we observed differences in various immune regulatory genes between bone marrow and cord tissue MSCs, especially for the “a” cluster cells (Figure 9C). Notably, BM-High_a MSCs had higher gene expression for PTGES2, and the protein encoded by this gene is known to have direct or indirect role in immunomodulation by MSCs(86, 87). PTGES2 encodes membrane-bound prostaglandin synthase E2 which converts prostaglandin H2 (PGH2) to prostaglandin E2 (PGE2) that is

known to have anti-inflammatory/immunosuppressive effects on various immune cells, including macrophages, T cells and B cells(88, 89).

MSC surface proteins are important for their significant roles in identification and functions(90). When we compared gene expression for surface markers that are known to have some immunomodulatory functions, BM derived MSCs showed higher expression for CD46, CD47, and CD276 whereas UCT derived MSC had higher expression for CD81. Surface expression of CD46 protein helps MSCs to inhibit complement binding and complement-mediated lysis(91). CD47 serves as a "don't eat me" signal to avoid phagocytosis by engaging its cognate ligand signal-regulatory-protein alpha (SIRP alpha) on phagocytes(92, 93), and the interaction of CD47 with SIRP alpha is reported to inhibit antigen presenting cell (APC) maturation and enhance STAT3 phosphorylation and IL10 induction in APC(94). CD276 is known to cause immune suppression by inhibiting T cell function and is currently being targeted as a check point blockade therapy for cancer(95); however, their specific role in MSC-mediated immunomodulation is not yet confirmed. CD81 is one of the surface markers used to identify MSC-derived extracellular vehicles (EVs) but does not have any reported immunomodulatory role for MSCs; however, CD81 coding gene is known to affect T regulatory (Treg) and myeloid-derived suppressor cell (MDSC) function enhancing tumor growth(96). Taken together, gene expression differences for surface markers related to immune response between BM and UCT-MSCs implicates potential differences in the immunomodulatory functions between BM and UCT-MSCs. Further, differences in immunomodulatory gene expression between the High_a and High_b clusters for both BM and UCT-MSCs indicates functional and phenotypic heterogeneity within each MSC product. No differences in expression of a

small number of pluripotent and stemness marker genes was detected between BM and UCT derived MSCs (Figure S6), though we note that abundance of these transcripts was very low which reduces power to observe differential expression.

In summary, this study both confirms the potential for functional differences to exist between MSCs derived from different tissues and even donors, and that within-sample heterogeneity is low. The expression of cell cycle markers is a major component of heterogeneity among donors, and manufacturing processes may need to accommodate biological and technical influences on proliferative potential. These findings will help improve the therapeutic MSC manufacturing processes and identify the most efficient cells from a heterogeneous MSC population. Even though differences in the gene expression profile between bone marrow and cord tissue G0 MSC were found, further studies are needed to confirm these results as well as the impact of these differences on the clinical use of these cells.

2.5. Experimental procedure

2.5.1. *Study approval*

This study was approved by the ethics committee of the institutional review boards at Georgia Institute of Technology and Duke University (IRB protocol no. H17348). All procedures involving human participants were in accordance with the ethical standards of the research committee. Informed consent was obtained from all participants.

2.5.2. *Human Bone Marrow MSC collection*

Seven frozen human bone marrow-derived MSC lots from six male donors were purchased from RoosterBio Inc., and expanded using RoosterBio expansion protocol (https://www.roosterbio.com/wp-content/uploads/2019/10/A.-RoosterBio-MSC-001-BOM-Expansion-Protocol-IF-08022016.pdf). Briefly, a BM-hMSC high performance media kit was brought to room temperature. Then 1 vial of Media Booster GTX (RoosterBio, catalog no. SU-003) was added to 500ml hMSC high performance basal media (RoosterBio, catalog no. SU-005). The 10million BM-hMSC vial was obtained from a liquid nitrogen dewar and immediately thawed at 370C for approximate 2 minutes while monitoring the process and removed from the water bath once a small bit of ice remained. Cells were aseptically transferred to a 15ml centrifuge tube and 10ml cultured media was added. The cells were spun down at 200g for 10 minutes and all the supernatants was discarded. The cell pellet was re-suspended in 10ml of culture media and transferred into 500ml media bottle. The cells were mixed by capping and gently inverting the bottle and distributed (seeded at 3500-4000 cells/cm2 and 42 mL media/T225) equally into T-225 vessels (Corning cat no. 431082). The vessels were transferred to a 370C incubator ensuring that the surfaces were covered with media. The vessels were observed microscopically from day 1 to determine percentage confluency. Once they reached >80% confluency, they were harvested the next day, and cryo-preserved in Cryostor CS-10 freezing media. All single-cell RNA-sequencing was performed on the out of thaw cells directly from these frozen vials.

Samples BM1 and BM2a were cultured in Laboratory A, while samples BM2b, BM3, BM4, BM5 and BM6 were cultured in Laboratory B. Samples BM2a and BM2b were from the same donor.

2.5.3. *Human Cord Tissue MSC collection*

For cord tissue derived samples, six frozen human MSC samples from four male donors were provided by the department of pediatrics, Duke University. Cryopreserved P0 vials were placed in a sterile bag which was itself placed in a 37°C water bath. Vials were thawed until the cell suspension was slushy (~2 minutes). Cell suspension from the vials were transferred to a 15 mL tube containing XSFM (Irvine Scientific, cat. no. 91149) using a sterile serological pipette and the cell suspension was mixed slowly. The cryovial was rinsed with 0.5 mL of XSFM and the rinse was transferred to the 15 mL tube. After mixing slowly, the cell count and viability was measured. The cells were mixed in the 15 mL conical using a sterile pipette and transfer the volume containing 3.4 x 106 cells into a HYPER flask containing 1.12 L of XSFM and the bottle was mixed gently. The HYPER Flasks were placed into a 37°C/5% CO2 incubator. The P1 cells were harvested after 5-7 days. The P2 cells were then frozen using CS-10 freezing media and cryopreserved. The Cryopreserved P2 cells were shipped to us for the downstream characterizations. All the single-cell RNA-sequencing was performed on the out of thaw cells directly from these frozen vials.

Samples UCT1a, UCT1b and UCT1c come from the same donor.

2.5.4. *Thawing and single cell suspension preparation for Single-Cell RNA-Sequencing*

Frozen vials containing 1 million MSC were thawed in a 370C water bath for a couple of minutes. Cells were then aseptically transferred to a 15 ml centrifuge tube. Room temperature RPMI media (1mL) was used to rinse the cell vial and added to the cells in the

15 ml tube. Another 3 ml of media was added to the cells and mixed well with serological pipette. Cells were counted using Nucleocounter and spun down at 200 g for 10 minutes. The cells were re-suspended in media and counted again and processed for scRNA-sequencing.

2.5.5. *Single-cell RNA-seq library preparation and sequencing*

The Illumina-Bio-Rad ddSEQ platform was used to process, capture, and barcode the cells to generate single-cell Gel Beads by following the manufacturer's protocol. Cell suspensions were loaded onto a ddSEQ Cartridge along with reverse transcription master mix, and encapsulated and barcoded by the Single-Cell Isolator. Lysis and barcoding took place in each droplet. Droplets were disrupted and cDNA was pooled for second strand synthesis. Libraries were generated with direct cDNA tagmentation using Nextera technology. Tagmentation was followed by 3′ enrichment and sample indexing to prepare indexed, sequencing-ready libraries. The libraries were sequenced using Nextseq sequencing Platform on an Ilumina NextSeq in the IBB Molecular Evolution core at Georgia Tech (PE75, mid-output V2.5 kit). The library quality (check for primer dimers, adopter dimers, ethanol contamination, degradation as well as the size and concentration) was confirmed before each sequencing run using Agilent Bioanalyzer2100.

All the BM-MSC samples have 2415 cells with an average 3,835 genes/cell (Table 1) and the UCT-MSCs have 1785 cells with average 3056 genes/cell (Table 2). The BM-MSC samples have average 28,890,505 (Table 1) reads per sample and the UCT-MSC samples have 33,771,805 (Table 2) reads per sample in average.

Confirmation that MSCs were relatively pure populations of undifferentiated cells was revealed by FACS analysis of cell surface markers provided by the manufacturer as a release criteria. Furthermore, scRNA-seq (which is less sensitive due to high drop-out rates) confirmed prevalent expression of NT5E (CD73), THY1 (CD90), and ENG (CD105) and absent expression of CD34 among other genes (Fig. 10). ENG was expressed on 66% of the cell, NT5E on 72%, and THY1 on 92%. In contrast, each of the transcripts PTPRC, CD34, CD14, ITGAM, CD79A, CD19, and HLA-DRA were detected in less than 0.5% of the cells.

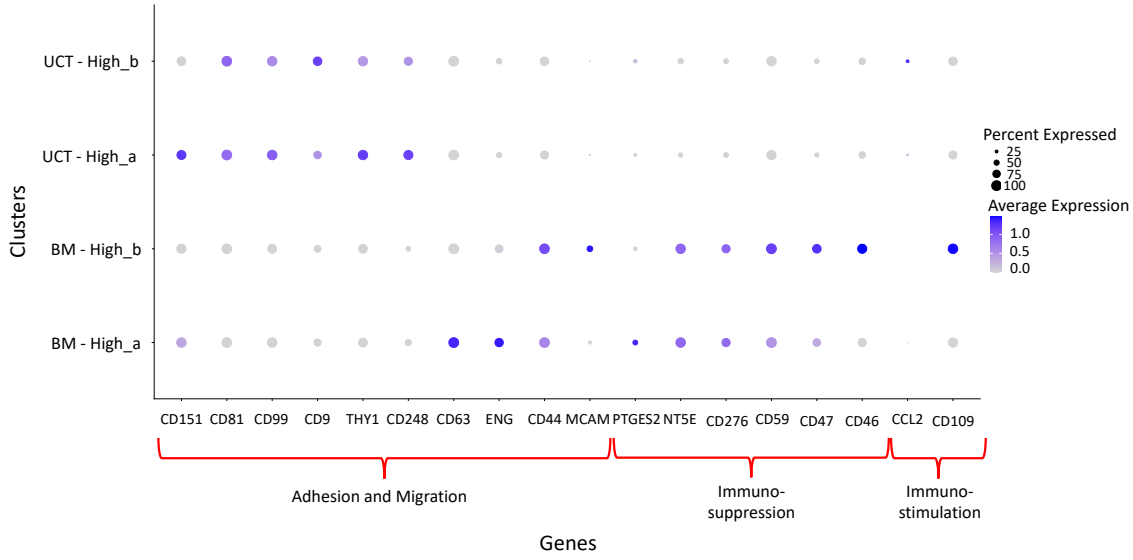


Figure 10. Dot plot displaying the average expression for cell adhesion and migration and immunomodulatory function-associated genes

The colors represent the average expression of the genes per cell. The scale is from low expression of 0 (gray) to high expression of 1.5 log counts per million (blue). The size of the dots represents the proportion of cells in each cluster that express the gene.

2.5.6. *Data Analysis*

Sample demultiplexing and gene counts were extracted using the Illumina SureCell pipeline. The raw reads were trimmed, and the gene-barcode matrix was generated.

SureCell was also used to filter and align the samples and to generate gene-cell UMI count matrices. The seven samples from bone marrow were sequenced in three different batches, and the six samples from cord tissue were sequenced in two batches.

Downstream analysis was initiated with SC3 software(83) for cell clustering. Owing to the high proportion of zero counts in most cells (as is typical of dropseq data), we elected to perform differential expression analysis on pools of pseudo-cells whose profiles have a distribution of read counts that closely resembles that of bulk RNAseq. Custom R scripts were used to generate pools of pseudo-cells by pooling groups of 20 cells within each sample and cluster, and then summing their gene count. The gene expression values from the pseudo cells were normalized to counts per million before using EdgeR(97) for normalization and differential expression estimation. Default likelihood ratio tests assuming negative binomial distributions were performed in EdgeR to evaluate the significance of differential expression. Ten permutations of this procedure were performed, and the average differential expression and negative-log P-values were computed, and genes significant with an FDR less than 5% were selected for downstream gene ontology analysis. Then gene ontology was performed on the differentially expressed genes using GSEA(98) and ToppGene(99) tools.

CHAPTER 3. CHARACTERIZATION OF THE EFFECT OF CYTOKINES ON ENHANCEMENT OF MESENCHYMAL STROMAL CELL FUNCTION

3.1. Abstract

Enhanced mesenchymal stromal cells are subject of multiple studies for their immunosuppressive capacities. Cytokines like TNFα or IFNγ are known to modify the anti-inflammatory effects of MSC. While previous studies used both TNFα and IFNγ as MSC immunomodulatory enhancers, it is not clear how these cytokines activate the immunosuppressive function of MSC. Previous in vitro studies have also shown suppression of lymphoid cells (B cells, NK and T cells), when they are exposed to enhanced MSC. The mechanism by which these cells suppress the immune system is yet to be clarified. Here I characterize the expression profiles of MSC treated with TNFα or IFNγ, using bulk RNA seq. 48 hours after exposure to the cytokines, MSC treated with IFNγ have higher levels of expression of immunosuppressive markers, even though compared to the control MSC, both treatments seem to activate similar MSC pathways. I also characterize the gene expression profile of T cells after exposure to IFNγ-treated MSC (γMSC) using single cell RNA seq. Suppression of the T cell mitotic cell cycle, as seen in previous studies, is confirmed, but differences in the response of two types of T cells (CD4 and CD8) to γMSC treatment are observed.

3.2. Introduction

When the body detects a foreign antigen, the immune system responds by activating a variety of immune cells(100). One of the first responder to the foreign antigen are the T cells. When activated, these T cells secrete cytokines that activate but also regulate the immune response(100, 101).

T cells are produced in the bone marrow and they are capable of recognizing a large variety of antigens. Some T cells recognize self-antigens, in which case they are generally inactivated on the thymus, the organ where T cell maturation takes place. This inactivation process is called tolerance induction. Tolerance usually ensures that T cells do not attack the cells that present self-proteins of the body. When tolerance induction fails, T cells that do recognize self-antigens are released and attack cells from the body. The malfunction of this process leads to different autoimmune diseases.

Similar self-damaging attacks of the immune system can also be seeing in Graft vs Host Disease (GvHD). In this case, the patients receive an organ transplant which contains donor's T cells. These T cells recognize the host cells as foreign and activate the immune system and attack the host cells. GvHD is the primary cause of morbidity and mortality after hematopoietic cell transplantation (HCT), afflicting as many as one half of all recipients(102).

Current treatments for many immune related diseases include the use of anti-inflammatory drugs and potent immunosuppressive and immunomodulatory agents(103). Despite their effect on immune responses, these agents are not uniformly effective and are

associated with substantial toxicities. They may also increase the risk of opportunistic infections or recurrent leukemia(104).

In recent years, researchers have contemplated the use of stem cells to treat autoimmune disorders. Even though cell therapy is already been used in non-responding patients (patients that do not respond to the standard treatments), the study of these cells is of increasing importance. Mesenchymal stem cells are known to have an inhibitory effect on immune cells including naïve, memory and activated T cells, B cell, NK cells and dendritic cells(105, 106). The immunosuppressive effect of the MSC have been shown to be enhanced by adding different kind of cytokines in the culture media(107-111).

There are two cytokines used in research to enhance the immunosuppressive activity of the MSC: IFNγ and TNFα(109, 111, 112). These two cytokines are pro-inflammatory cytokines, secreted by activated T cells. Even though it is known that enhanced MSC suppress the activity of some immune cells and can also induce apoptosis, the interaction between MSC and the immune cells has not yet been defined in terms of gene expression(105, 113-115). This suppression can be done by different pathways such as the interferon mediated suppression(116) and TNF signaling(117).

One of the main problems when trying to analyze the effect of MSC on the immune system is the reproductivity of the findings. As described in Chapter 2, MSCs can come from different tissues and donors, which impacts their therapeutic potential. In addition to tissue and donor related variability, these cells may be treated with different cytokines for a period of time between 48 to 72 hours, to enhance their immunomodulatory capacities, creating a cytokine and time related variability.

This Chapter aims to assess the effect of different enhancement treatments on bone marrow derived MSC gene expression profiles, and subsequently the effect of the enhanced MSC on the gene expression profiles of mouse T cells. To accomplish this, I characterize the gene expression profiles of bone marrow derived MSCs treated with IFNγ and TNFα. I then describe the gene expression profiles of these cells at different times after the exposure to cytokines (T0, T48 and T72). Finally, I characterize the gene expression profile of mouse T cells after exposure to enhanced bone marrow derived MSC.

3.3. Results

3.3.1. *Comparison of the gene expression profile of Mesenchymal Stem Cells from different donors following exposure to IFNγ and TNFα*

A total of 12 samples from 4 different donors were collected: each donor has 3 samples, one cultured in normal media with no cytokines added, one in media + IFNγ and one in media + TNFα. These samples were then sequenced using bulk RNA-seq. The first 5 principal components (PC) capture 88% of the variance in gene expression, and as show in Figure 12A,B treatment is the major variance component (62.4% weighted contribution, mostly due to PC1 and PC2), but donor is also an important variant component as well (36.1%, mostly PC3-5). The PCA plot in figure 11C, shows that IFNγ treated samples have the most extreme response, as they separate from the control and TNFα treated samples along PC1, whereas TNFα treatment induces a greater change along PC2. Donor effects are also observed, donor, with Donors 1 and 2 having more negative scores, and donors 3 and 4 more positive (toward the top of the plot) regardless of the control or treatment status.

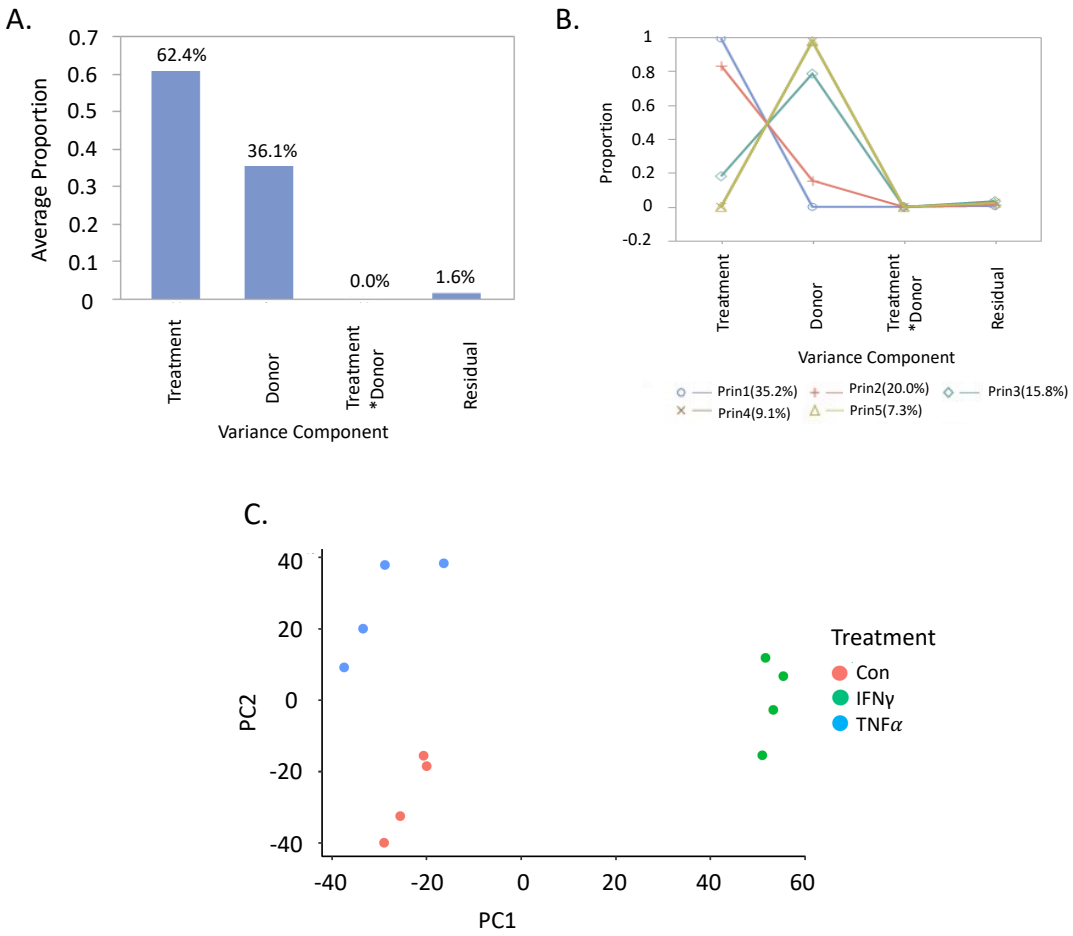


Figure 11. Principal variance component analysis and principal and Principal component analysis.

(A) Treatment is the major source of variability within the MSC samples, followed by donor. (B) the first 5 principal components capture 88% of the total variability. Principal components 1 and 2 represent mostly the treatment variability, while principal components 3, 4 and 5 represent the donor variability. (C) PC1 clearly separates the IFNγ treated samples (green) from the TNFα (blue) and control (salmon) samples.

Differential expression analysis contrasting the IFNγ treated and control samples revealed that 3,561 genes were significantly differentially expressed at the 5% False Discovery Rate: 1,885 genes were overexpressed in IFNγ treated samples and 1,676 genes were overexpressed in control samples. On the other hand, only 1,223 genes were

significantly differentially expressed between TNFα treated samples and control samples: 739 genes were overexpressed in TNFα treated samples and 484 genes were overexpressed in control samples.

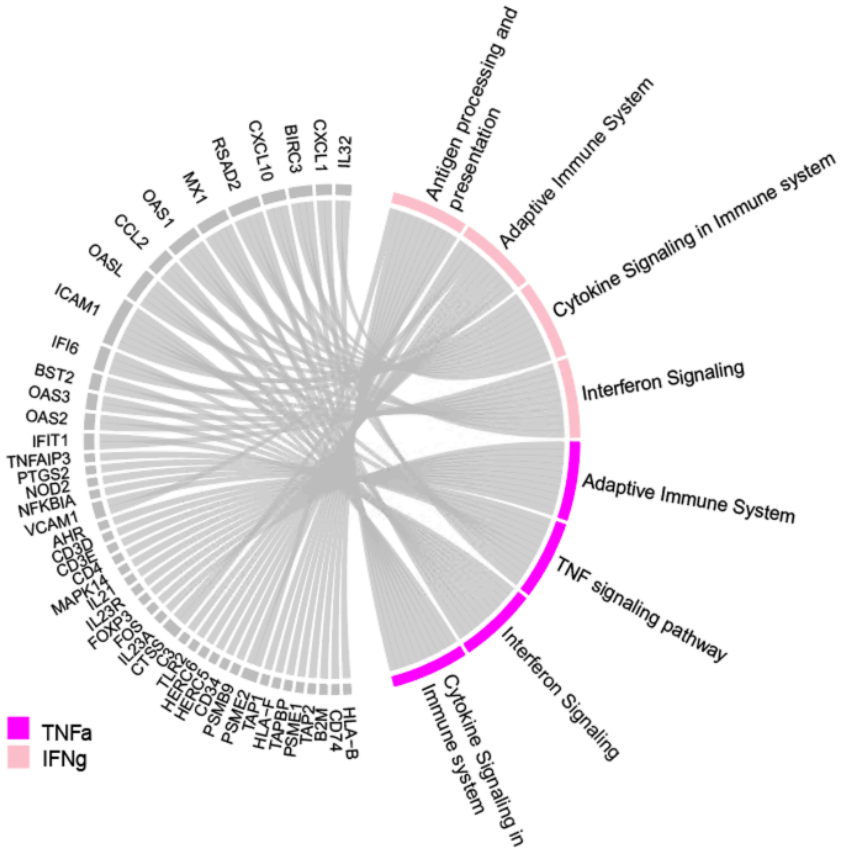


Figure 12. Differentially expressed genes and pathways in IFNγ and TNFα treated samples, compared to the control samples

Chord diagram summarizing differential expression and enriched pathways in the IFNγ (pink) and TNFα (purple) treated samples, compared to the control samples.

Gene ontology analysis revealed that similar pathways were overexpressed in both types of treated samples compared to the controls. Among the pathways overexpressed in the treated samples, interferon signaling, antigen processing and presentation and cytokine signaling in the immune system stand out (Figure 12). Each ark links a gene indicated on

the left to a pathway on the right, and in almost half the cases statistically significant over-expression is seen the Interferon, TNF signaling, and adaptive immune pathways after both treatments. The treated samples thus overexpress genes and pathways related to the suppression of the immune system.

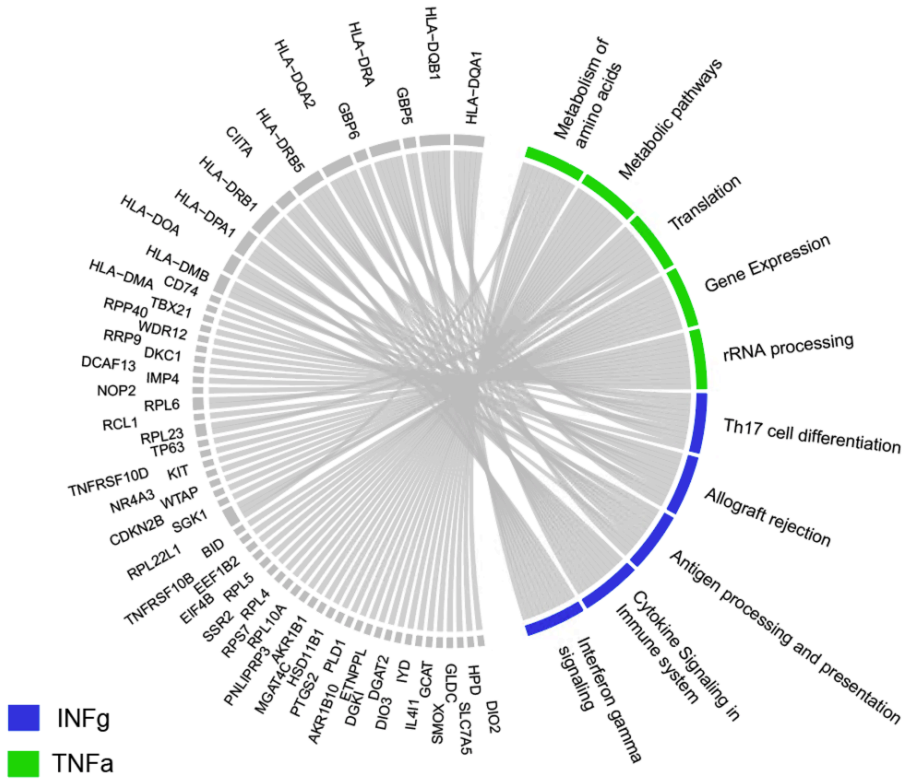


Figure 13. Chord diagram displaying some of the enriched pathways in IFNγ treated samples and TNFα treated samples, when compare to each other.

I also performed differential expression analysis between IFNγ and TNFα treated samples. 4,796 genes were found to be significantly differentially expressed, with 2,397 gene overexpressed in IFNγ treated samples and 2,399 genes overexpressed in TNFα treated samples. The gene ontology analysis performed in these genes (Figure 13) shows an enrichment in pathways related to gene expression, translation and metabolism on the

samples treated with TNFα, while the samples treated with IFNγ overexpressed pathways related to the repression of the immune system.

Table 3. Immunomodulatory markers.

PTGES2	ENG	CD63	CD44	CD46	CD47
NT5E	CD276	CD151	CD99	THY1	CD248
IDO2	NOS1	NOS2	CCL2	CD151	MCAM
CD109	CD59	IDO1	PTGS2	CD81	CD9

The differential expression of immunomodulation markers is summarized in Table 3 and Figure 14. Each of these markers is up-regulated following either cytokine treatment, but consistently the level of overexpression of the immunomodulation markers is higher when the samples are treated with IFNγ than with TNFα. Consequently, the remainder of the study focused on MSC treated with IFNγ.

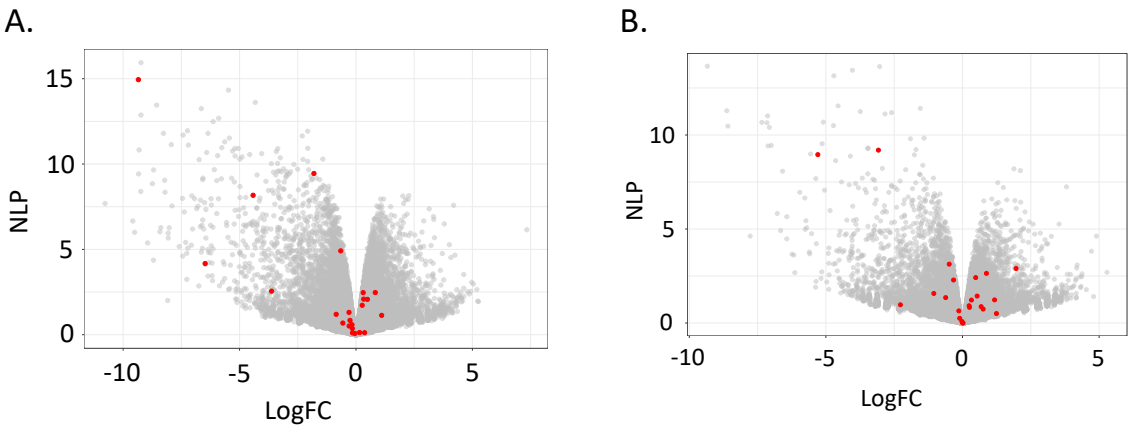


Figure 14. Volcano plots of the differentially expressed genes between IFNγ and TNFα treated samples and control samples.

(A)Volcano plot of the differentially expressed genes between IFNγ treated MSC and control samples. (B)Volcano plot of the differentially expressed genes between TNFα treated MSC and control cells. In both plots, the dots in red represent the immunomodulatory markers, showing a higher level of overexpression in the samples treated with IFNγ.

3.3.2. *Evaluate the gene expression profile of (γMSC) through time: at T0, T48 and T72*

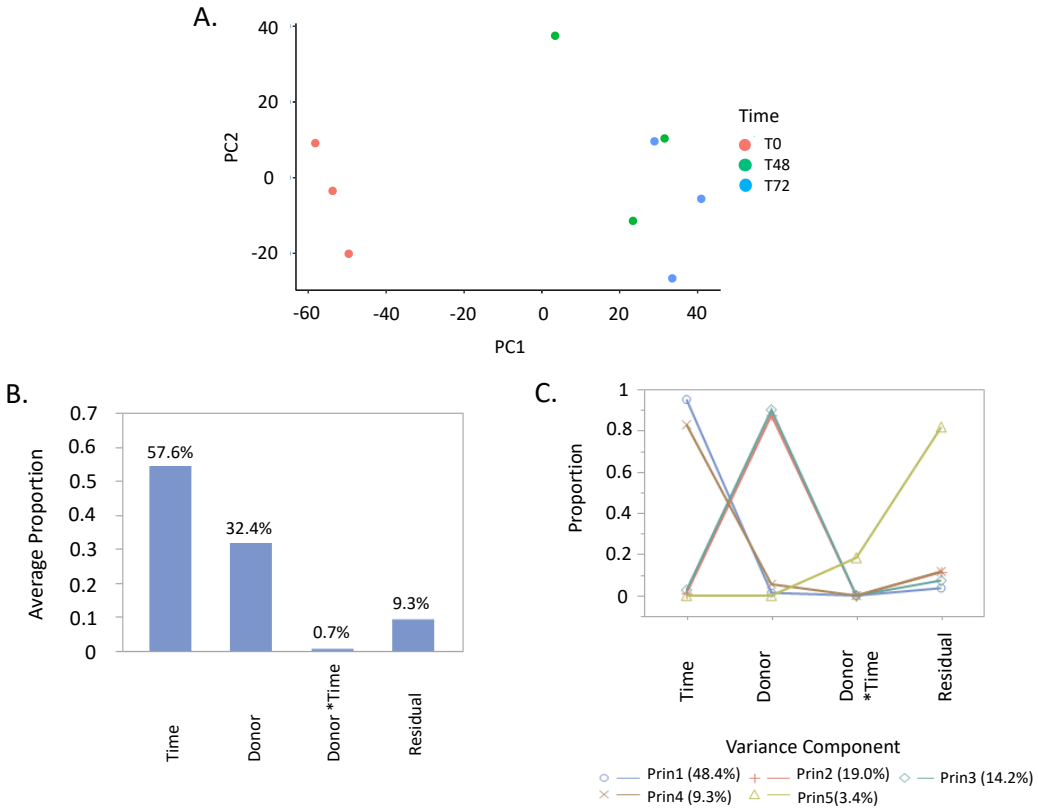


Figure 15. Principal variance component analysis and principal component analysis.

(A) PC1 plots separated the samples by treated and control, while PC2 separates the samples by donor. (B) The major source of variability in the gene expression profile of these cells is the time (0h, 48h and 72h), followed by the donor variability. (C) Time variation is represented by PC1 and PC4, while donor variation is represented by PC2 and PC3. PC4 represents a combination between donor and time variance.

In order to evaluate the time course of expression profiles of (γMSC) through time (at T0, T48 and T72 hours), 9 bone marrow derived MSC samples from 3 different donors were sequenced using Bulk-RNA: each donor has a control sample and two treated samples, one sequenced 48 hours after exposure to IFNγ and one sequenced 72 hours after

exposure to IFNγ (Figure 15). The first 5 principal component account for 94% of the total variance. Principal component 1 and 4 represent the time related variance (57.6%), while principal components 2 and 3 represent the donor related variance (32.4%).

2,372 genes were differentially expressed between the control samples at T0 and the treated samples sequenced 48 hours after exposure to IFNγ, at a FDR lower than 0.05: 1,077 genes downregulated and 1,295 upregulated at T48, including JAK2, STAT1, STAT2, STAT3, MYD88, IRAK4, IRF3, IRF9, IRF7 and many HLA genes, consistent with expectations of a strong interferon response.

Very similar results were obtained contrasting the 72hr treatment with baseline, but only 34 genes were observed to be significantly differentially expressed between the T48 and T72 samples. There was a slight trend toward enhanced differential expression at 72 hrs, but no obvious induction or repression of new pathways. This result shows that the interferon response is consistent for up to three days of treatment.

Gene set enrichment analysis detected differential expression of genes involved in TGFβ signaling, reactive oxygen species generation, KRAS signaling, cholesterol homeostasis and bile acid metabolism, as well as the inflammatory and Interferon gamma responses.

3.3.3. *Effect of γMSC on T cells*

I next asked what impact cytokine treatment has on the ability of MSC to modify the gene activity of T cells *in vitro*. Murine T-cells were cultured with or without γMSC conditioned media, resulting in 6 scRNA-seq samples for comparison: 3 representing

activated T cells in different cell cycle phases (TG0, TG1, TSG2) and 3 representing suppressed T cells, also in different cell cycle phases (G0γMSC, TG1γMSC, TSG2γMSC). Each cell cycle phase was captured by flow sorting based on …

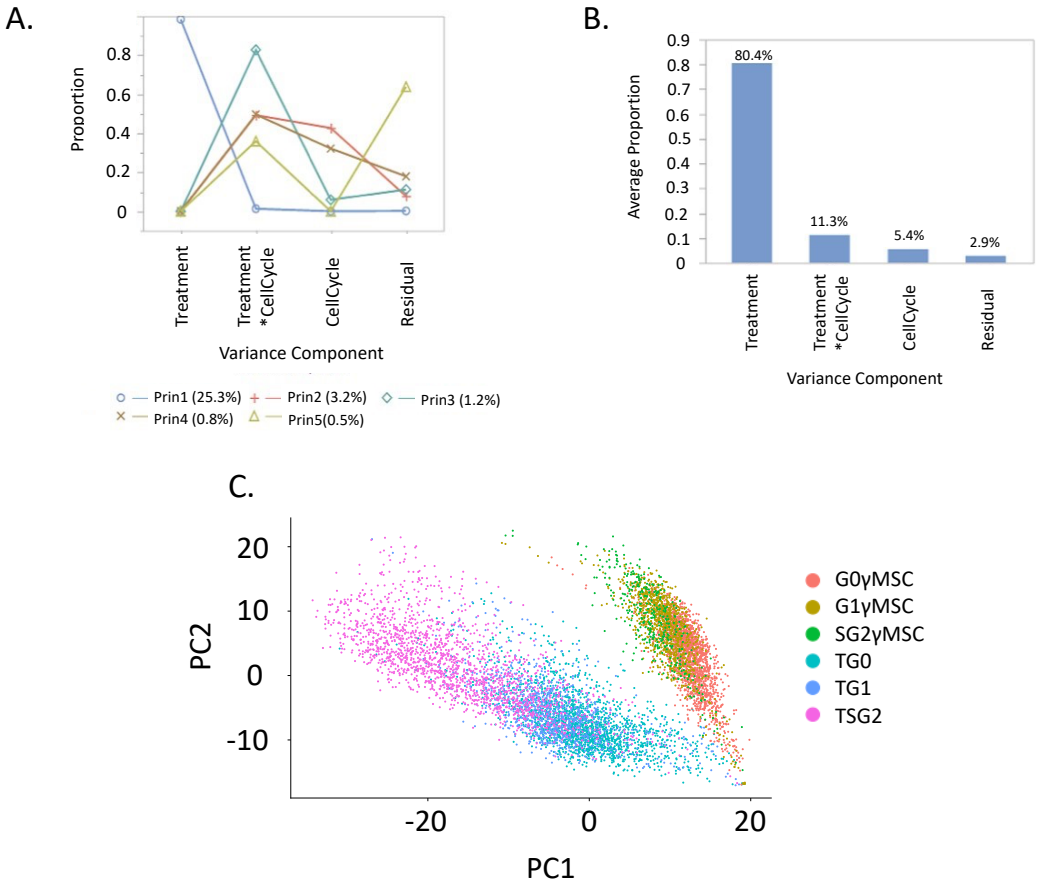


Figure 16. Principal variance component analysis and principal component analysis.

(A) Treatment variation is represented by PC1 and PC4, while donor variation is represented by PC2 and PC3. PC4 represents a combination between donor and time variance PC1 plots separated the samples by treated and control, while PC2 separates the samples by donor. (B) The major source of variability in the gene expression profile of these cells is the time (0h, 48h and 72h), followed by the donor variability. (C).

PCA was again performed on the single cell data, as summarized in Fig. 16A and 16C. The variance related to the treatment correspond to the 80.4% of the total variance

due to the first 5 PC, while the cell cycle corresponds to just 5.4% (Fig.16B). PC1 separates the samples mostly by treatment, on the right of PC1 are the samples treated with yMSC, and on the left, the control samples. On the other hand, PC2 separates the samples by cell cycle phase and treatment (Fig.16A). High PC2 corresponds to cells in S, G2 and M phases, low PC2 correspond to the cells in G0 phase. Cells in G1 phase are located in between.

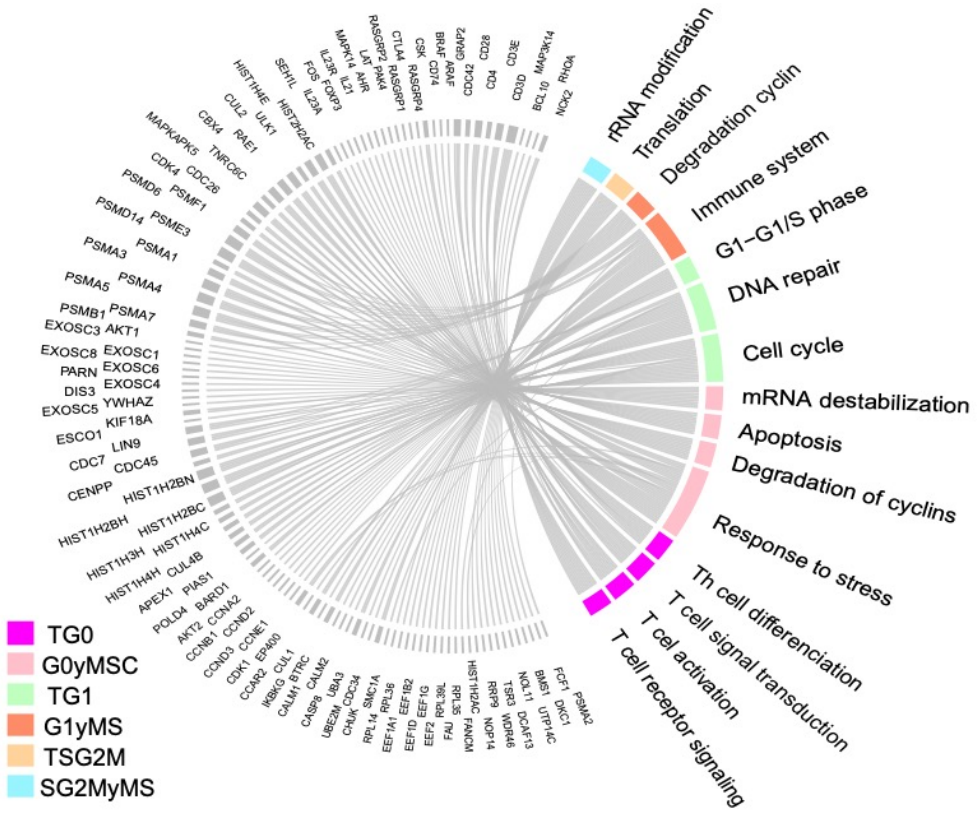


Figure 17. Chord diagram of the enriched pathways from the treated vs control samples.

Differential expression and gene ontology analysis performed on the significantly differentially expressed genes (FDR < 5%) resulted in the following findings:

TG0 vs G0yMSC

7678 genes were differentially expressed between the G0 T cells treated with untreated and interferon-gamma exposed MSC. The activated T cells in G0 phase express genes involved in T cell activation, T cell receptor signalling pathway, T cell signal transduction and Th cell differentiation, among others. On the other hand, the T cells in G0 suppressed with yMSC treatment expressed genes involved in cellular responses to stress, degradation of cyclins, apoptosis and mRNA destabilization (Figure 17).

TG1 vs G1yMSC

7769 genes were found to be differentially expressed. The activated T cells in G1 phase express genes involved in translation, cell cycle, DNA repair and mitotic G1-G1/S phase. The T cells in G1 suppressed with yMSC and the T cells in G0 suppressed with yMSC have a similar gene expression profile. The T cells in G1 suppressed with yMSC overexpress genes involved in cellular response to stress, degradation of different cyclins but also innate immune system genes (Figure 17).

TSG2M vs SG2MyMSC

7390 genes were found to be differentially expressed in the mitotically active T cells. The activated T cells in S, G2 and M phases express genes involved in the cell cycle, DNA repair and translation, while the T cells in S, G2 and M phases, suppressed with yMSC, express genes involved in cellular response to stress, rRNA modification, innate immune system signaling and MHC mediated antigen signaling processing and presentation (Fig. 17).

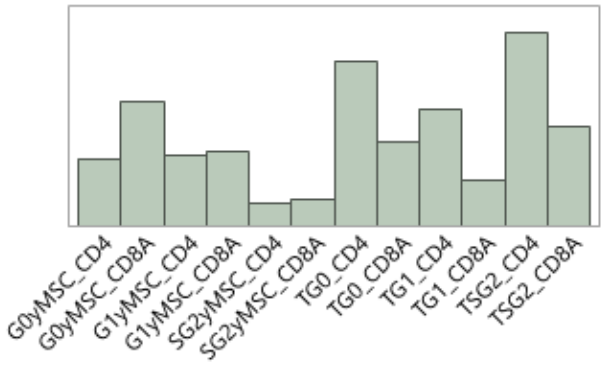


Figure 18. Abundance of CD4 and CD8A T cells per sample.

In a final analysis, cell markers (CD4, CD8A, CD8B) were used to differentiate $CD4^+$ and $CD8^+$ T cells (Table S1). $CD4^+$ cells predominate with about a 2:1 ratio to $CD8A^+$ in all cycle phases without treatment. However, after the T cells are treated with yMSC the global ratio of CD8 to CD4 positive cells is altered toward parity, implying a significant reduction in the proportion of $CD4^+$ relative to $CD8^+$ T cells. Notably as well, yMSC treatment reduces dramatically the proportion of cells in S/G2 (Figure 18), consistent with suppression of T-cell division.

3.4. Discussion

MSC are currently used for cell therapy in different kinds of diseases. Due to their immunomodulatory capacities, these cells are being tested in different diseases related to the immune system(115, 118-121). Even though these cells are currently used in cell therapy, it is important to characterize them and the effect they have on the immune system.

The immunomodulatory capacities of MSC in vitro can be enhanced by exposure to immunosuppressive cytokines(109, 111, 112). INFγ and TNFα are both use to enhance the

immunomodulatory capacities of the MSC, making them an interest asset for immune diseases.

Comparing bone marrow derived MSC samples, before and after exposure to INFγ and TNFα., both treatments resulted in overexpression of immunomodulatory markers, like PTGES2, ENG, NT5E, IDO1 and IDO2, and enrichment of pathways related to the immune system, like cytokine signaling and antigen processing and presentation, compare to the control samples. Even though these treated samples overexpressing similar genes and pathways compare to the control samples, the PCA analysis show they have different gene expression profiles, with the TNFα treated samples being placed near the control samples. Since the PCA analysis suggested there is a difference between the INFγ and TNFα treated samples, I examined more closely the expression of various immune suppressive markers in the different samples. Once again, the level of overexpression of these markers when using INFγ was greater than the levels of overexpression of the same markers when using TNFα, relative to control samples. Two possible explanations are that there could be a delay in the time course of full activation of the TNFα treated MSC, or that the activation of these genes in the TNFα treated samples never reaches the levels observed in INFγ treated samples. Since the difference in treatment effects was observed in four different donor MSC samples, it is unlikely that it reflects inter-individual differences in the expression of a key receptor, but rather is intrinsic to the wiring of MSC to receive different cytokine signals.

Another potential source of variability when using MSC to suppress the immune system, is the culture time. Previous studies have determined the activation time of these

cells to be 48 hours after being expose to the cytokines, recommending these cells to be used at this time(109, 122). Unfortunately, this is not always the case. MSC are cultured in conditioned media for a period of 48 to 72 hours before these are used for research or therapy. This procedure variability raises the question of whether the activation period of these cells affects their function: are the MSC still active after 72 hours, or do they have to be used exactly 48 hours after the exposure to the cytokines. Our results suggest that there is little effect of the additional 24 hours incubation, since the gene expression profiles of the samples sequenced after 48 hours compare to the samples sequenced after 72 hours of exposure to INFγ were very similar. Only 34 genes were found to be significantly differentially expressed at an FDR lower than 0.05, suggesting the function of these cells remain the same from 48 to 72 hours.

The third experiment in this study focused on the effect of INFγ enhancement of MSC, with respect to modulation of co-cultured T cells. Previous studies(123-126) have shown there is suppression of the T cell mitotic cycle upon exposure to cytokines, which led us to prepare samples at separate cell cycle stages (G0, G1 and S/G2/M) using flow cytometry. These cells were then separated into two groups, INFγ treated and control cells. For this analysis, single cell RNA seq was used to characterize the gene expression profile of each cell individually, rather than bulk RNAseq, in order to consider whether there is heterogeneity in the cell-to-cell responses. The single cell expression analysis confirmed the suppression of the cell cycle by the enhanced MSC and also showing an increase on the number of cells in the G0 phase on the treated samples. Contrary to previous studies(111), INFγ treated MSC don't seem to suppress the T cell activation. Furthermore, expression of the cell markers CD4 and CD8A showed an important decrease in the

abundance of $CD4^+$ T cells in the treated samples compare to control sample, suggesting the γMSC promote differentiation of T-cells away from the CD4 helper profile toward the more cytotoxic CD8 state.

In summary, this study showed there is a difference on the gene expression profile of the MSC when they are activated with INFγ or TNFα, suggesting a difference in the activation of the immunomodulation capacities related to the cytokine. When INFγ is used to activate the MSC immunosuppression capacity, there is no significant difference on the expression profile of these cells between 48 and 72 hours after exposure to the cytokine. Finally, I demonstrated that γMSC have the capacity to suppress cell cycle progression of T cells, and also to influence the balance of helper and cytotoxic T cell types. These in vitro observed properties need to be reassessed in vivo by examining the expression profiles of resident T-cells after MSC therapy.

3.5. Methods

3.5.1. *Data Analysis Bulk RNA*

Quality control of the short sequencing reads was performed using FastQC version 0.11.5 (127) and the alignment of the fastq files to either mouse (mm10) or human reference genomes (hg38) was performed using the STAR aligner version 2.5.2a (128), using quantMode function to extract the gene count matrix.

The bulk RNA seq analysis was done using EdgeR version 3.26.8 (97). The gene expression values from the samples were normalized to counts per million and differential

expression was assessed using the default likelihood ratio tests assuming negative binomial distributions.

3.5.2. *Data Analysis of single cell RNA*

Sample demultiplexing and gene counts were extracted using the Illumina Sure cell pipeline. The raw reads were trimmed, and the gene-barcode matrix was generated. SureCell was also used to filter and align the samples and to generate gene-cell UMI count matrices.

Downstream analysis was initiated with Seurat software version 3.2.0(82) for normalization and PCA. Owing to the high proportion of zero counts in most cells (as is typical of dropseq data), we elected to perform differential expression analysis on pools of pseudo-cells whose profiles have a distribution of read counts that closely resembles that of bulk RNAseq. Custom R scripts were used to generate pools of pseudo-cells by pooling groups of 20 cells within each sample and cluster, and then summing their gene count. The gene expression values from the pseudo cells were normalized to counts per million before using EdgeR version 3.26.8 (41) for normalization and differential expression estimation. Default likelihood ratio tests assuming negative binomial distributions were performed in EdgeR to evaluate the significance of differential expression. Ten permutations of this procedure were performed, and the average differential expression and negative-log P-values were computed, and genes significant with an FDR less than 5% were selected for downstream gene ontology analysis. Then gene ontology was performed on the differentially expressed genes using GSEA(98) and ToppGene(99) tools.

CHAPTER 4. COMPARISON OF BONE MARROW ASPIRATE CONCENTRATE FROM HEALTHY AND OSTEOARTHRITIC DONORS USING SINGLE CELL RNA-SEQ

4.1. Abstract

Osteoarthritis is degenerative joint condition affecting more than 30 million adults in the US. Even though this is a common disease, the is not cure for it. The current treatment is based on lowering the levels of pain and inflammation with the use of steroids and ibuprofen. Physical therapy or surgery are used in patients nonresponding to the anti-inflammatory drugs. Recent studies have showing promising results in the use of cell therapy as a treatment for osteoarthritis. Bone marrow aspirate concentrates (BMAC) are used as treatment for osteoarthritis. They contain a low percent of mesenchymal stem cells, which are multipotent cells with immunosuppressive capacities. The BMAC used normally autologous, which means they come from the patient itself. In this study, I compare the gene expression profile of the different immune cells contain on the BMAC samples from osteoarthritic (OA) patients and non-osteoarthritic (non-OA) donors, showing differences in the abundance of the cell types, like NK, T cells, mature B and monocytes, among others, which seems to be due to the individual. Beside the abundance variation, I showed a difference in the gene expression profile of the cell depending on the donor condition, OA and non-OA. The OA samples show an enrichment in genes involved in inflammatory processes compare non-OA samples.

4.2. Introduction

Osteoarthritis is a common degenerative joint condition, usually occurring in the knee, hip, hands or spine, which currently affects over 32.5 million adults in the US, according to the CDC website (https://www.cdc.gov/arthritis/basics/osteoarthritis.htm). To date, there are no licensed disease-modifying osteoarthritis drugs (DMOADs), capable of inhibiting the structural changes associated with the disease. Treatment plans involve medication and exercise regimes to reduce pain and improve joint function, but eventually provide no options for total recovery. There is thus considerable interest and investment in use of cell therapy to circumscribe these rather drastic interventions.

The osteoarthritic joint typically develops after ischemic or mechanical damage to the subchondral bone and articular cartilage and is characterized by a progressive breakdown of cartilage, the development of subchondral bone sclerosis and osteophytes, and hypertrophy and inflammation of the synovial membrane. Since the inflammation is considered low grade, this condition has been described as “non-inflammatory” arthritis. The degradation of cartilage, and the subsequent phagocytosis of the fragments by the synovium, render the synovium hypertrophic and activated. Inflammatory cytokines and proteolytic enzymes further degrade the cartilage matrix, and this is exacerbated by the involvement of T cells, B cells, monocytes and macrophages(129-132).

Bone marrow aspirate concentrate (BMAC) is used in clinical trials to improve symptomatic osteoarthritis pain and joint function(133-135). The BMAC can come from the patient itself, autologous, or from a donor, allogeneic. Since bone marrow aspirate only

contains 0.001-0.01% mesenchymal stromal cells (MSC) of the nucleated cell fraction(136), it requires centrifugation to concentrate cells, growth factors and platelets. The resulting BMAC is a more concentrated source of lymphocytes (13%), eosinophils (2.2%), monocytes (1.3%), basophils (0.1%) platelets, and MSC (0.03%) (137, 138).

The factors responsible for BMAC's activity are unknown, but thought to be related to platelets, their associated growth factors and the MSC fraction. It has been proposed that the elderly population have a lower MSC count, with reduced proliferative and chondrogenic capacity(139). BMAC isolates with higher cell concentration have been suggested to result in better treatment outcomes(140) and it has also been proposed that patients with mild OA, Kellgren-Lawrence grade 2 (KL2) experience better outcomes with BMAC treatment than those with advanced OA (KL4)(47, 141). BMAC administration has been reported to minimize fibrocartilage formation. Monocytes, and their progeny, osteoclasts and macrophages, are thought to drive the pathogenesis of OA and so reducing their activation could be important in preventing the progression of the condition. Anti-inflammatory mediators such as VEGF, IL-1RA, and PDGF, derived from platelets, could invoke a response to bring regeneration in the joint. BMAC is often used as an adjunct therapy, and given its variability patient to patient, this complicates any proof of efficacy. While BMAC samples from the patient itself are one of the options for cell therapy to treat OA in non-responding patients, it is crucial to understand the different in function between OA and non-OA derived BMAC.

Single cell RNAseq has become a very valuable tool to investigate the cell populations present and the gene expression levels of tissue samples. In order to understand the composition of BMAC and its attributes which may contribute to a potential mode of

action, particularly for easing of joint pain, single cell RNAseq has been used to analyze BMAC samples in this study.

This study aimed to understand the BMAC composition and gene expression profile variability due to the donors and the donor's condition (OA or non-OA). A single cell RNAseq library of BMAC from 21 healthy control individuals(142) was used to compare with single cell RNAseq analysis of BMAC isolated from osteoarthritic patients enrolled in an osteoarthritis trial (the MILES study). OA grade, age and BMI are co-factors considered in the comparison, along with activation state of monocytes and pro-inflammatory/anti-inflammatory mediators.

4.3. Results

1.1.1. *Evaluation of the content of bone marrow aspirate concentrates from osteoarthritis patients compared to bone marrow aspirate concentrates from healthy donors.*

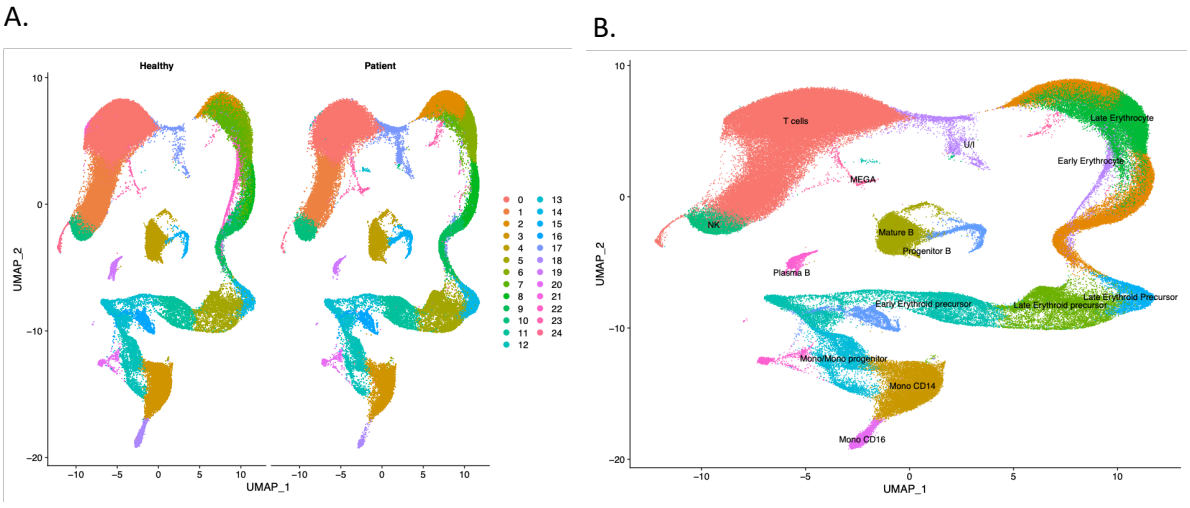


Figure 19. UMAP displaying the clusters and cell types present in the BMAC samples from OA and non-OA.

For this analysis I had 21 samples from 21 OA patients, and 25 samples from 19 non-OA donors(142), sequenced using single cell RNA sequencing. The average read depths of both datasets were similar with x0,000 reads and y,000 uniquely expressed transcripts per cell. The samples clustered in 25 groups. These subgroups were then assigned likely cell-type identities using published cell markers(142). All the cell types were present in both OA and non-OA datasets (Fig. 19). None of the subgroups correspond to MSC cells, though this is not surprising due to the expected low abundance of MSCs in BMAC.

Since this study focuses on the immune related cells and the anti/pro-inflammatory mediators, the erythroid cells were excluded from the remaining analysis. The data was re-clustered with Seurat, and the cell types were again identified using published gene markers (142)(10.1172/jci.insight.124928). In this case, we were able to identify all the previous cell types (without the erythroid cells) and a MSC cluster (Fig. 20A). Additional cell markers were also used to identify the MSC cluster (Fig. 20B).

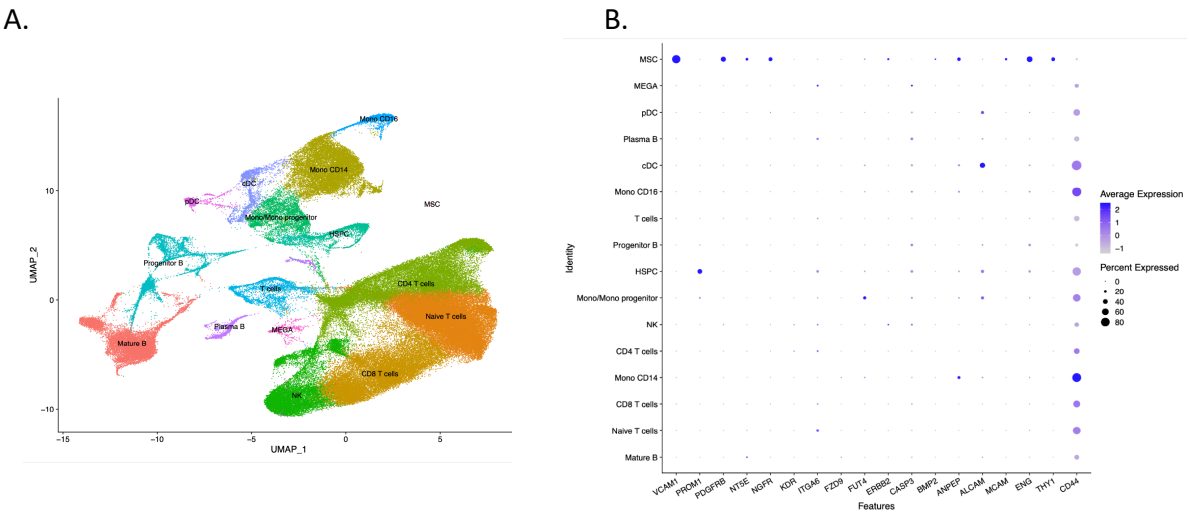


Figure 20. UMAP of the cell identification after removal of erythrocytes.

Figure 21A shows the content of the BMAC is not the same between OA and non-OA donors, as there is high variability of the cell ratios between the two conditions. A principal component variance analysis was performed to elucidate the major sources of variation between samples within conditions. As expected, cell type was the major source of variation (57.6% of the total variance), followed by the donor at just (3.7% of the total variance), BMI, sex, age and KL score (Figure 21B), suggesting the variance between samples is mainly due to the donor and the condition. An anova test was done to analyse the variation of the cell abundance depending on the sex, BMI, age and KL score. Some cell types such as monocytes, NK, progenitor B and MEGA shown a significant variation on the cell abundance related to the KL score and sex (Figures S10 and S11).

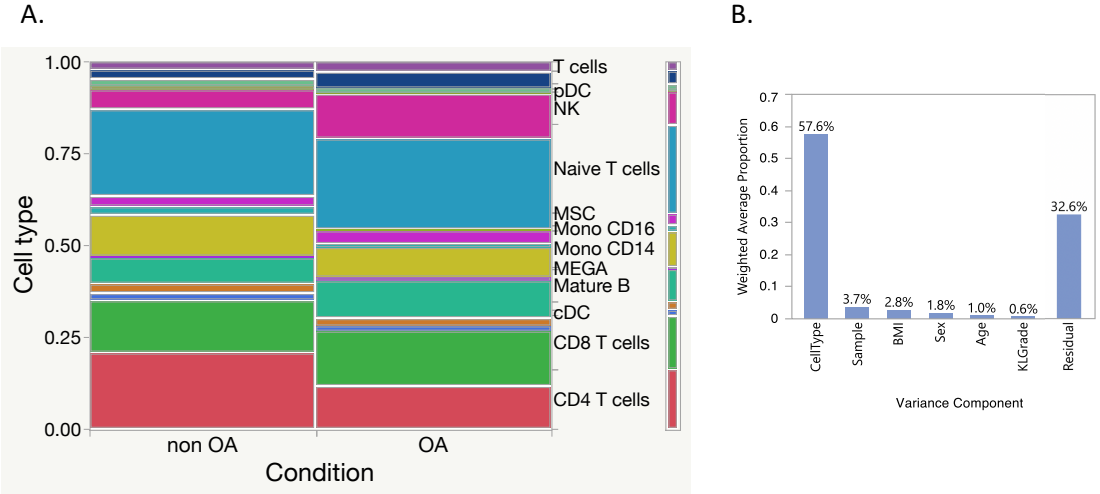


Figure 21. Sources of sample variability.

(A) Cell type distribution between conditions. (B) Principal component variance analysis.

4.3.1. *OA versus non-OA gene expression profiles*

Differential expression analysis was done between the OA and non-OA sample for each cell type. This analysis was done using the Wilcoxon rank sum with the default

parameters from Seurat, using FDR 5% as the significance threshold. I was not able to perform a DE analysis on the MSC due to their low count.

Table 4. Differential expression analysis between OA and non-OA BMAC.

Cell type	T cells	NK	MEGA	Plasma B	Progenitor B	Mature B	Mono and mono progenitor	Mono CD14	Mono CD16
Number of DE genes	844	1,054	773	665	1,472	993	1,098	1,301	1,324
Genes overexpressed in OA	613	770	240	374	1,032	726	817	989	1,040
Genes overexpressed in non-OA	231	284	533	291	437	267	281	311	284

Pathway analysis shows overexpression of genes and pathways related to immune system activation, in the OA samples compare to the non-OA samples (Figure 22).

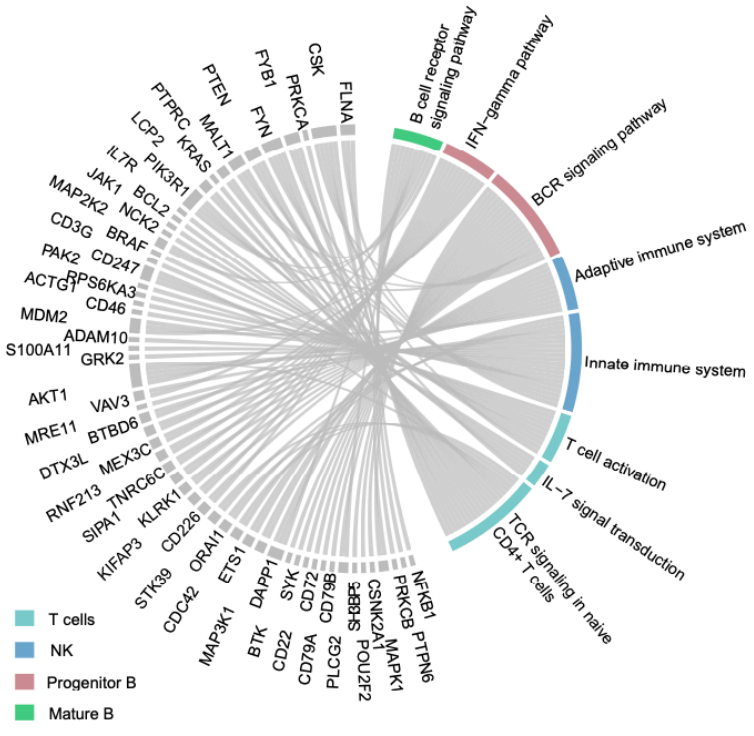


Figure 22. Chord diagram displaying the genes overexpressed in some of the cell types from OA derived BMAC compare to non-OA BMAC.

4.4. Discussion

Osteoarthritis is a common degenerative joint condition with no cure. The current treatment aims to reduce the joint pain and improve its function. Cell therapy is becoming a popular treatment for patients with OA(143). BMAC is often used as an adjunct therapy, but given its variability patient to patient, it is complicated to proof its efficacy. It is then critical to have a clear understanding of the function of cells used for this purpose and their variability. BMAC contain multiple cells types like T cells, monocytes, B cells and MSC. It's not clear whether the patients' outcome is due to a mixture of contributions from all the cell types included in the BMAC or if it's only due to the effect of the MSC immunosuppression capacity enhanced by the pro inflammatory cytokines present in the OA patient.

My preliminary analysis shows that the main source of variability in the BMAC gene expression profiles, beside the cell types variability, is due to the cell type, follow by the donor (sample) variability (Figure 21B). Sex, BMI and age seem to have a lower effect on the cell variability. The differential expression analysis further shows significant differences on the gene expression profiles of the cell types present in the BMAC related to the presence or absence of osteoarthritis. The gene ontology analysis performed in the significant differentially expressed genes suggest the cell types from the OA patients tend to have a higher activity of immune related functions such as TCR signaling, BCR signaling and T cell activation.

Previous studies have shown the role of the different immune cell types on the progression of the OA by secreting catabolic cytokines and stimulating protease destruction

of cartilage matrix(144, 145). The secretion of pro inflammatory cytokines by immune cells is an indicator of the disease status and progress. Since BMAC from OA patients have a higher expression of immune related pathways and pro inflammatory cytokine, it may be helpful to consider the use of BMAC form healthy donors as treatment for osteoarthritis. This would necessarily entail allogeneic specimens, which may however raise concerns over donor-mediated immune reactions.

4.5. Experimental procedure

1.1.1. *Single cell RNA-seq and data pre-processing*

Cryopreserved samples were thawed in 37°C water bath for ~2 minutes. A washing step was performed immediately after diluting the samples in 10ml PBS+0.1% BSA. The cells were counted for each sample using AOPI dye on Cellometer (Nexcelom) and Nucleocounter (Cemometec) automated cell counters. The samples were then processed through a magnetic removal of dead cells using the Dead Cell clean-up kit (Miltenyi Biotek, cat. no. 130-090-101) according to the manufacturer's protocol.

scRNA-Seq was performed using 10X Genomics Single Cell 3′ Solution, version 3.1, according to the manufacturer's instructions (protocol rev C). Libraries were sequenced on Nextseq500 (Illumina) and Novaseq6000 (Illumina). The resulting data were analyzed using Cell Ranger version 3.1.0 (10X Genomics). The OA samples were sequenced in 6 batches.

4.5.1. *Data Analysis*

Quality control metrics were used to select cells with a mitochondrial gene percentage less than 50% and at least 200 genes detected. The data were then analysed using Seurat's Louvain clustering algorithm, with a resolution of 0.4. The visualization was done using UMAP(146). We used Seurat Integration pipeline(82) to integrate the OA and non-OA datasets. This method aims to identify shared cell states that are present across different datasets, even if they were collected from different individuals, experimental conditions, technologies, or even species. It focuses on the identification of 'anchors' between pairs of datasets. These represent pairwise correspondences between cells from each dataset. A list of markers(142) was used to identify the different cell types present in the BMAC. A differential expression analysis between OA and no OA, for each cell type, was done using the Wilcoxon rank sum test from Seurat. We used ToppGene for the gene ontology analysis. The principal component variation analysis included cell type, donor, BMI, sex, age and KL score as variance sources and was performed on the first 5 PCs.

CHAPTER 5. CONCLUSION

The immune system is a complex system that protects the body against diseases. It detects a variety of pathogens, as well as cancer cells, by differentiating them from the organism's own healthy tissue. Malfunction of the immune system can cause autoimmune diseases, inflammatory diseases or cancer. There are more than 80 different autoimmune diseases(147), over 100 different cancers(148) and many more diseases that involve the immune system affecting people around the world.

Treatment for immune related diseases varies going from Acetaminophen(149) to decrease the pain, to chemotherapy(150), to kill fast growing cells. For many of these diseases, such as autoimmune diseases and graft versus host disease, the gold standard treatment are not always effective, leading to an important unmet medical need.

Cell therapy has become a target for research as possible treatment for these diseases. Multiple cells are used for cell therapy such as mesenchymal stem cells, hematopoietic stem cells and bone marrow aspirate concentrates. These cells can also come from different tissues and they can be extracted from the patients itself or from a donor. Even though the studies about cell therapy show promising results(151-157),these sources of variabilities across the field confound the ability to compare clinical trial results and contributed to a lack of conclusive data to support their potential for clinical use(4). It is then crucial to provide a characterization of diverse batches from various tissue sources in order to help evaluate reasons for success or failure of individual trials or patient responses.

The second chapter of this thesis focuses on the study of the MSC transcriptomic profile variability related to the tissue of origin and the donor, using scRAN-seq technology. This analysis shows the cells derived from the same tissue cluster in two group, a group of cells in G0 phase and a group of cells with high expression of mitotic genes. This difference seems to be related to the donor. Previous studies(5, 6, 76-79) have also show the cell cycle gene expression as a major source of heterogeneity. According to the Huang et. al., 2019(76) study, the cell cycle is related to the immune regulatory potency of the MSCs, reason why I did not regress out this sources of variance before performing downstream analysis as done in previous work(77). I also demonstrate that bone marrow and umbilical cord tissue derived MSC have different transcriptomic profiles. The umbilical cord tissues derived MSCs show a higher expression of mitotic genes compare to bone marrow derived MSC. The G0 cells have different transcriptional properties that are directly relevant to their biological functions such as immunomodulatory potential. These differences likely explain some of the documented differences between them. The donor differences having an impact on the gene expression profile of these cells can also contribute to the difference seen in clinical trials.

On the third chapter of this thesis, I focused on the immunomodulatory capacities of the bone marrow derived MSC and their effect on the T cells gene expression profile. The immunomodulatory capacities of MSC in vitro can be enhanced by exposure to immunosuppressive cytokines such as INFγ and TNFα(109, 111, 112) creating a source of variability. The comparison of bone marrow derived MSC samples, before and after exposure to INFγ and TNFα, using bulk RNA-seq, shows an overexpression of immunomodulatory markers, like PTGES2, ENG, NT5E, IDO1 and IDO2, and enrichment

of pathways related to the immune system, like cytokine signaling and antigen processing and presentation, in the treated samples compare to the control samples. Even though these treated samples overexpress similar genes and pathways compare to the control samples, a different in their gene expression profile was seen (Figure 11C). I examined more closely the expression of various immune suppressive markers in the different samples. The level of overexpression of these markers when using INFγ was greater than the levels of overexpression of the same markers when using TNFα, relative to control samples (Figure 14). This could be explained by a possible delay in the time course of full activation of the TNFα treated MSC, or a difference in the level of activation of these genes in the TNFα treated samples.

Another potential source of variability when using MSC to suppress the immune system, is the culture time. Due to technical circumstances, the MSC are cultured in conditioned media for a period of 48 to 72 hours before these are used for research or therapy. This procedure raises the question of whether the activation period of these cells affects their function. Our transcriptomic results suggest the additional 24 hours incubation have little effect on the genes expression profile of the MSC, with only 34 genes being found to be significantly differentially expressed at an FDR lower than 0.05, suggesting the function of these cells remain the same from 48 to 72 hours.

The third experiment in this chapter focused on the effect γMSC on the T cells transcriptomic profile, sequenced with scRNA-seq technology. The results confirm the suppression of the T cell mitotic cycle upon exposure to γMSC, as seen in previous studies(123-126). Contrary to previous studies(111), INFγ treated MSC don't seem to suppress the T cell activation. The expression of the cell markers CD4 and CD8A showed

an important decrease in the abundance of $CD4^+$ T cells in the treated samples compare to control sample, suggesting the γMSC promote differentiation of T-cells away from the CD4 helper profile toward the more cytotoxic CD8 state.

For the four chapter of this thesis, I analyzed the single cell gene expression profile of bone marrow aspirates, currently use as treatments for osteoarthritis. As mentioned before, there are different sources of variability that can have an impact on the cell's gene expression profile and thus in the patient's outcome. In this case, I compare the content and gene expression profile of 21 BMAC samples from osteoarthritic patients and 25 sample from non-osteoarthritic donors.

My preliminary analysis shows that the main source of variability in the BMAC gene expression profiles, beside the cell types variability, is due to the cell type variability, follow by the donor (sample) variability (Figure 21B). Sex, BMI and age seem to have a lower effect on the transcriptomics profile. The differential expression analysis further shows significant differences on the gene expression profiles of the cell types present in the BMAC related to the presence or absence of osteoarthritis, showing a higher activity of immune related functions such as TCR signaling, BCR signaling and T cell activation (Figure 22).

Since Previous studies have demonstrated the involvement of the immune system on the progression of the OA through the secretion of catabolic cytokines (144, 145) it is important to assess the activity of the immune system cells on the BMAC samples. These results suggest it may be helpful to consider the use of BMAC form healthy donors as treatment for osteoarthritis.

In summary, this thesis shows transcriptomic analysis are useful to assess the difference in single cells and whole tissues, gene expression profiles. We could prove there are multiple sources of variability when using cells as possible treatments for different diseases. These sources of variability can be related to donor itself, his health condition or tissue of origin, and should be taking into account previous to their clinical use. A deeper characterization of these cells needs to be done in order to have more concise results during the clinical trials and a better understanding of their role in the patient's outcome.

APPENDIX A. SUPPLEMENTARY INFORMATION CHAPTER 2

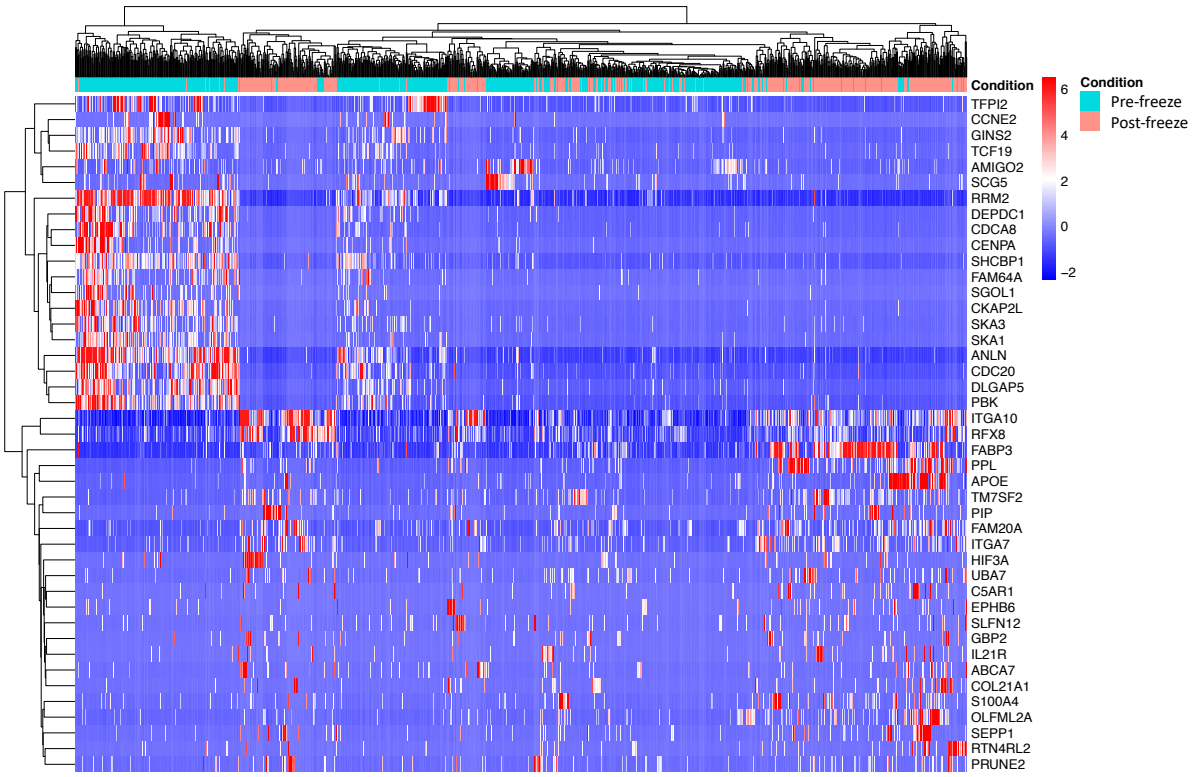


Figure S1: Heatmap displaying the top 50 DE genes for the between post-thaw and pre-freeze MSC comparison. The DEG analysis between these samples shows a significant overexpression of 1,743 genes on the pre-freeze samples, compared to 310 genes significantly overexpressed in the post-thaw samples. The most significant overexpressed pathways on the Pre-freeze samples are cell proliferation and cell adhesion, while the pathways over-expressed in the frozen samples are cholesterol/Steroid biosynthesis and cell death regulation.

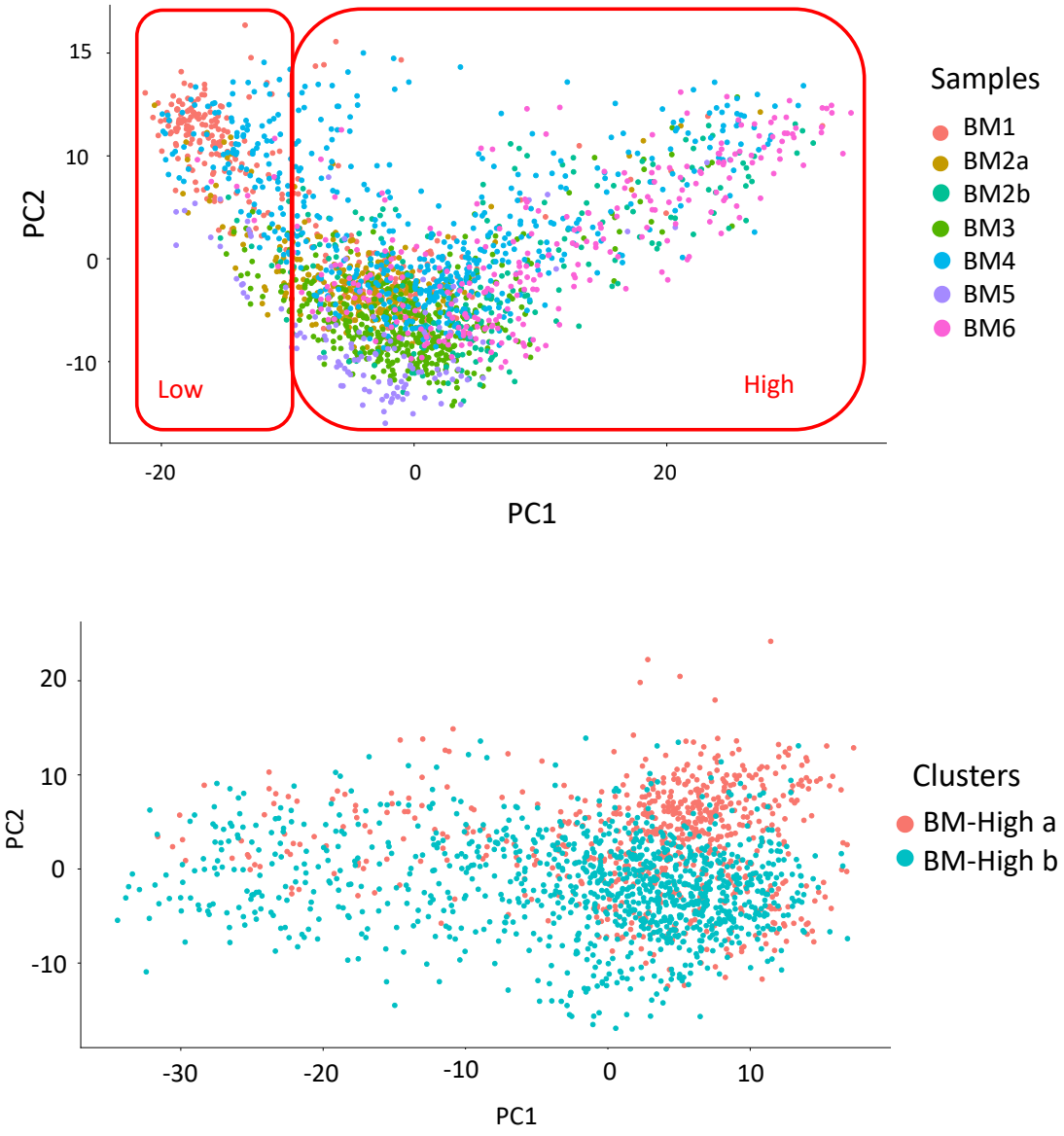


Figure S2: PCAs created with Seurat (A) Two major clusters differentiate along PC1. These clusters correspond to the low (left) and high (right) UMI count cells. (B) This PCA shows the clusters BM-High_a and BM-High_b. This cluster are not as well differentiated when using Seurat as they are when clustering the cells with SC3.

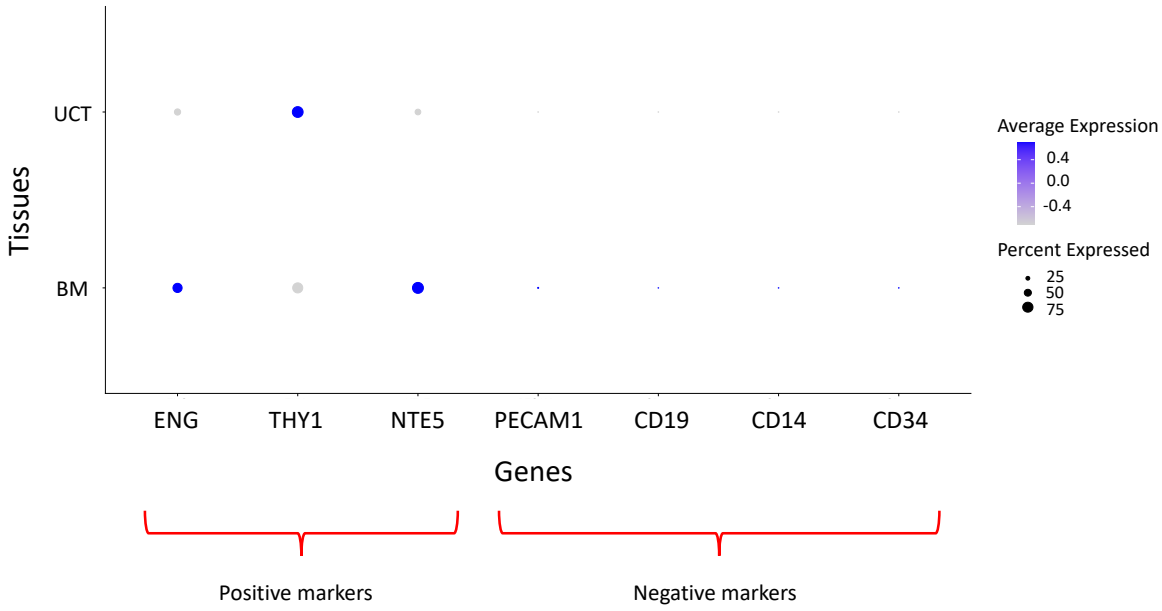


Figure S3: Dot plot displaying the MSC identity markers established by the ISCT. The size of the dots corresponds to the percentage of cells, in the tissue, expressing the gene. The color corresponds to the average non-zero expression of the gene, per cell, in each tissue. Light purple represents low expression per cell, while dark purple corresponds to high expression per cell.

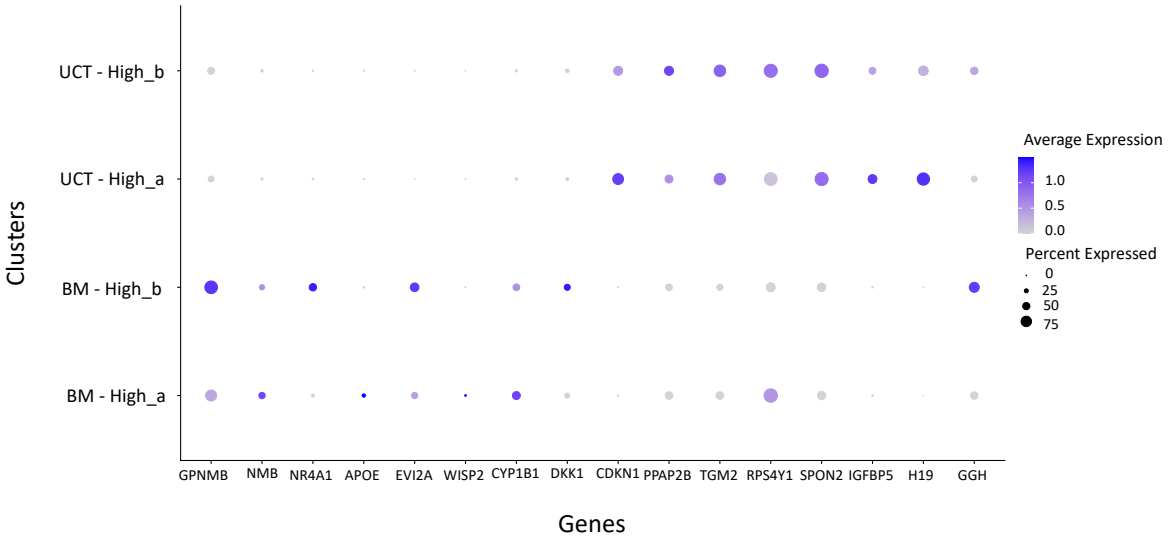


Figure S4: Dot plots displaying the top 16 differentially expressed genes in each group. The first 8 genes are overexpressed in BM derived MSC while the last 8 genes are overexpressed in UCT derived MSCs.

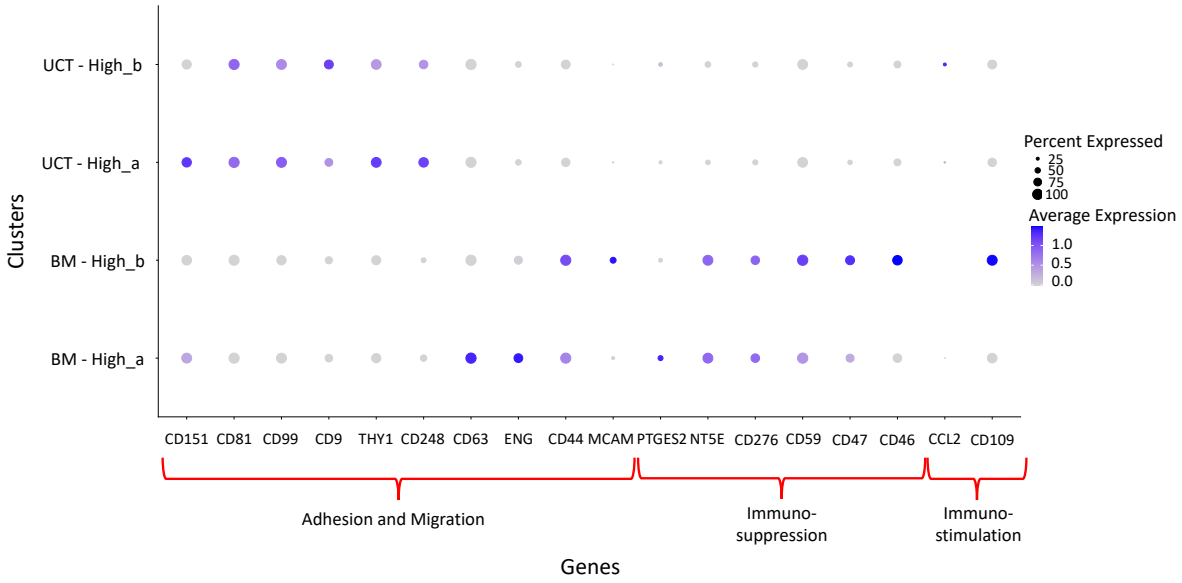
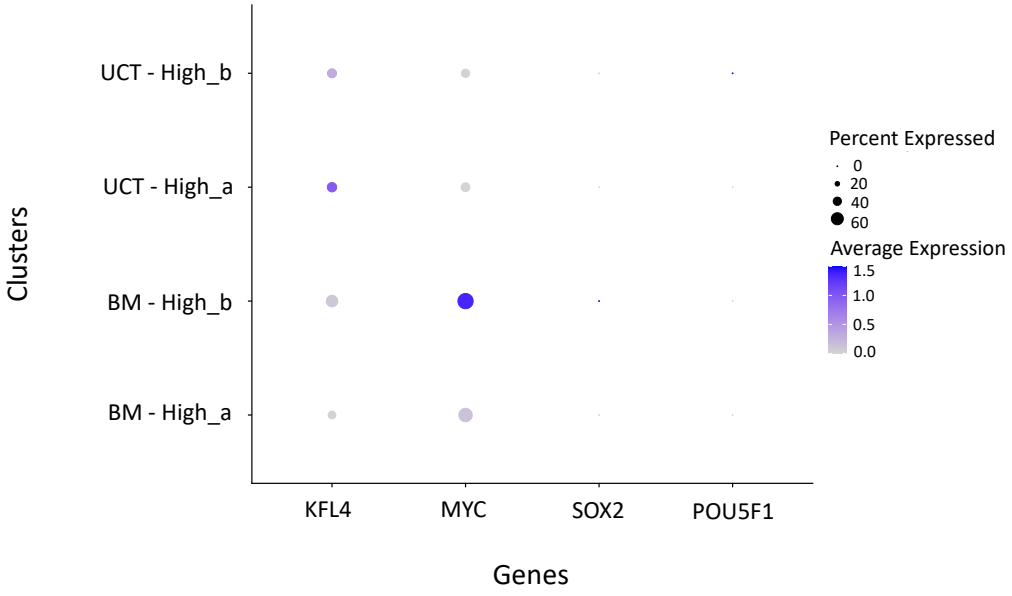


Figure S5: Dot plot displaying the average expression for cell adhesion and migration and immunomodulatory function-associated genes. The colors represent the average expression of the genes per cell. The scale is from low expression of 0 (gray) to high expression of 1.5 log counts per million (blue). The size of the dots represents the proportion of cells in each cluster that express the gene.

A.



B.

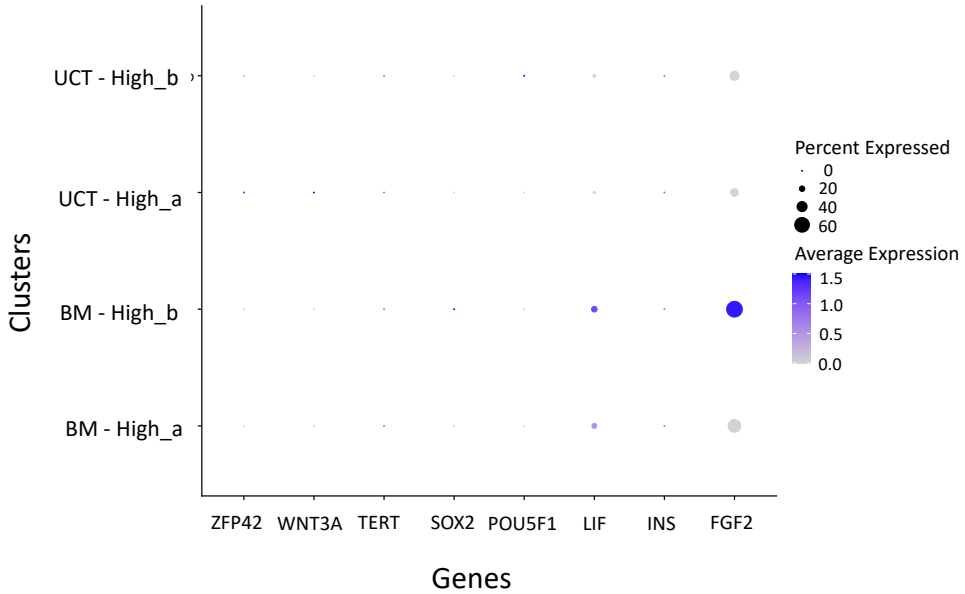
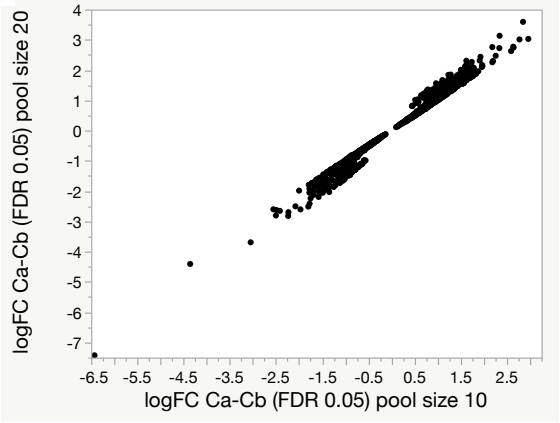


Figure S6: Dot plots displaying genes of interest. (A) This dot plot shows some pluripotent markers. (B) This dot plot displays stemness markers. The expression of some of these genes is low. There is not significant difference in the levels of expression between the clusters of MSCs.

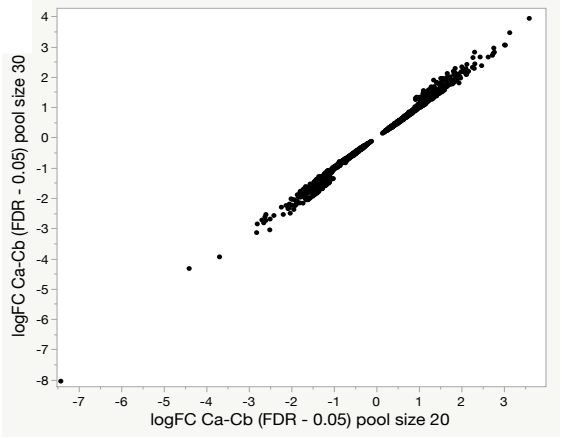
A.

	Pool size 10	Pool size 20	Pool size 30
Genes after filtering	16,881	16,868	16,811
Significant DE genes	5,155 (30.5%)	4,639 (27.5%)	4,043 (24.5%)
Up regulated genes	2,524	2,355	2,049
Down regulated genes	2,631	2,284	1,994

B.



C.



D.

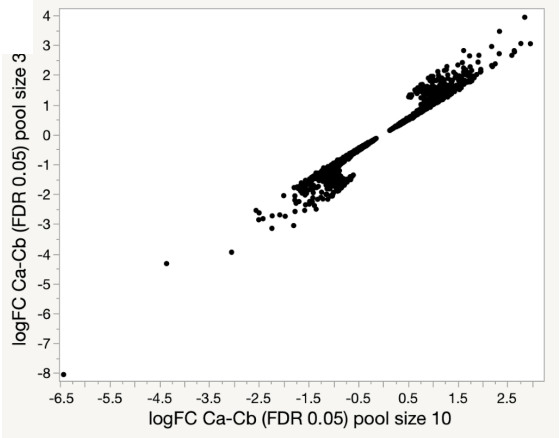
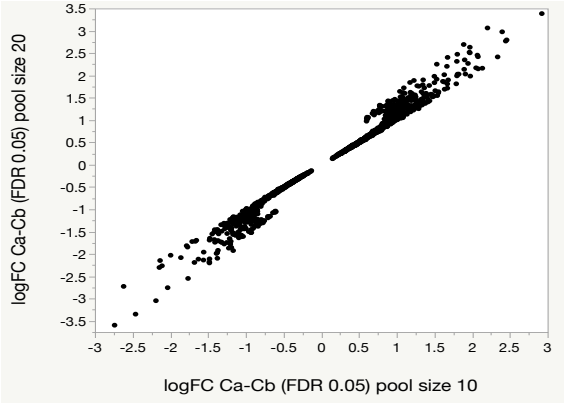


Figure S7: Differential expression analysis between BM-High_a and BM-High_b using scPool with different pseudocells sizes. (A) This table shows the number of significant differentially expressed genes found with scPool using pool sizes of 10, 20 and 30 cells. (B) This plot displays the log fold change of the significant differentially expressed genes found with a pool size of 10 compare to the significant differentially expressed genes found with a pool size of 20. (C) This plot displays the log fold change of the significant differentially expressed genes found with a pool size of 10 compare to the significant differentially expressed genes found with a pool size of 30. (D) This plot displays the log fold change of the significant differentially expressed genes found with a pool size of 20 compare to the significant differentially expressed genes found with a pool size of 30.

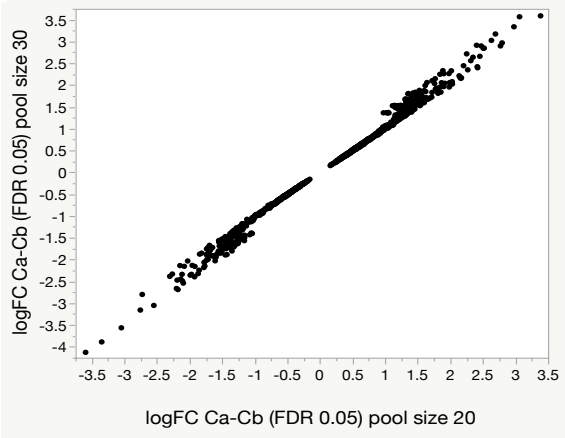
A.

	Pool size 10	Pool size 20	Pool size 30
Genes after filtering	16,742	16,711	16,704
Significant DE genes	2,549 (15.2%)	2,094 (12.5%)	1,878 (11.3%)
Up regulated genes	1,341	1,086	953
Down regulated genes	1,208	1,008	925

B.



C.



D.

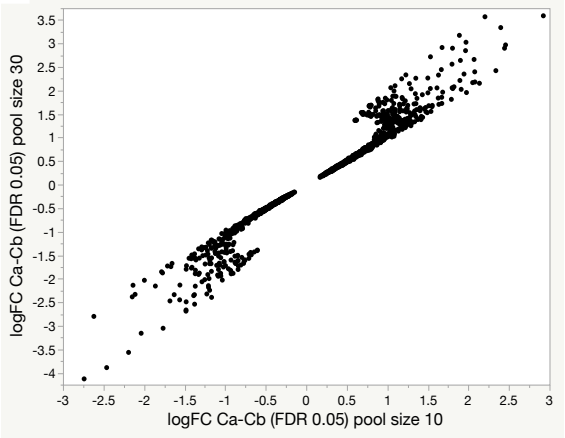


Figure S8: Differential expression analysis between UCT-High_a and UCT-High_b using scPool with different pseudocells sizes. (A) This table shows the number of significant differentially expressed genes found with scPool using pool sizes of 10, 20 and 30 cells. (B) This plot displays the log fold change of the significant differentially expressed genes found with a pool size of 10 compare to the significant differentially expressed genes found with a pool size of 20. (C) This plot displays the log fold change of the significant differentially expressed genes found with a pool size of 10 compare to the significant differentially expressed genes found with a pool size of 30. (D) This plot displays the log fold change of the significant differentially expressed genes found with a pool size of 20 compare to the significant differentially expressed genes found with a pool size of 30.

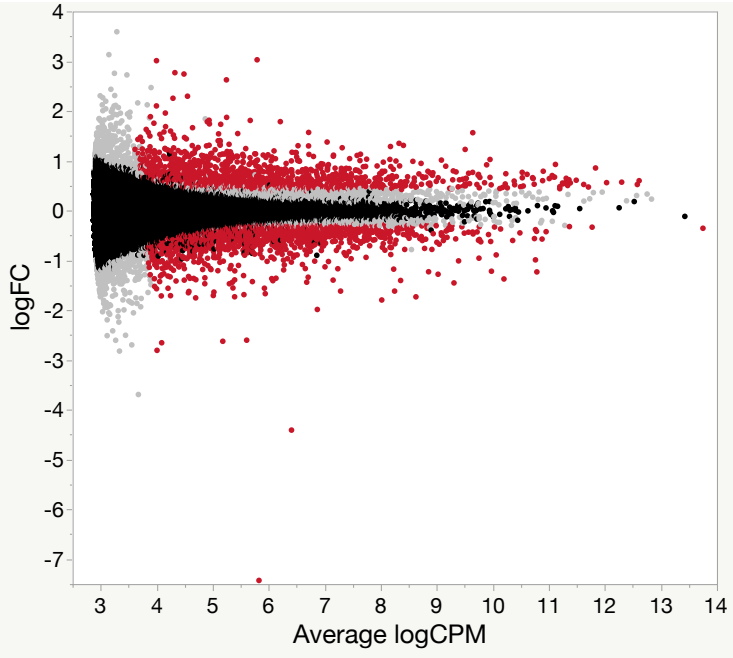


Figure S9: MA plot of the differentially expressed genes between BM-High_a and BM-High_b. The significant differentially expressed genes found using scPool, with a pool size of 20, and the Wilcoxon rank test in Seurat, are colored in red. In grey are the significant differentially expressed genes found with scPool, with a pool size of 20, but were not differentially expressed when using the Wilcoxon rank test from Seurat.

APPENDIX B. SUPPLEMENTARY INFORMATION CHAPTER 3

Table S1: CD4 and CD8 markers used to separate the T cells single cell dataset

CD4 markers	CD8 markers
CD4	ROBO1
EPB41L3	XCL1 /// XCL2
MAFB	KLRB1
TIAM1	HOMER1
CFP	LAMB3
CPVL	FANCF
LTF	NT5E
PTK2	GZMK
MARCKS	ELOVL4
C19orf6	RIT2
PTGER4	SLC14A2
KLF4	C8orf4
HK3	EFHC1
F5	SLC38A4
FAM13A	UGT2A3
PLXNB2	CDKL3
CDKN1C	CFH /// CFHR1
EPB41	TNIP3
IRAK3	CTSW
GPR22	CCR5
LILRB2	KLRC1 /// KLRC2
SERPINA1	KLRK1
C11orf9	STBD1
SOX4	KLRC3
CSF1R	PLA2G16
TNFRSF11A	KTELC1

GTF2A1L /// STON1-GTF2A1L	CRTAM
IL13RA1	KLRC4
ESR1	TNFSF14
SLCO1B1	MAP9
ANKRD55	ZNF167
FRMPD4	CD8B
COL14A1	CD8A
NOVA1	MEGF8
ADRA1A	MGC2889
IL8	ZNF80
ANK3	CYorf15B
TPM2	SLAMF7
GPT	ESRP1
MSR1	CCL5
LILRB3	C1orf129
RGS12	PTH2R
NR4A1	TLR3
LPHN2	AGAP1
POU5F1P4	ENDOG
COL5A2	IL7
RTN1	GJC1
GPR109B	GPR87
CORO1B	GPR162 /// LEPREL2
C11orf80	C17orf60
OR1D4 /// OR1D5	SLC27A6
VCY /// VCY1B	TARP /// TRGC2
PFKFB1	DHTKD1
LILRA2	NPVF
RAB7A	SLC25A21
GRB14	ADAM3A
TCF7L2	CAMTA1
PRSS1	RPRM

RUNX1T1	SCGB1D1
TMEM176B	TMOD2
GRM8	RPS4Y1
TNFRSF4	RABL5
MGC31957	GZMA
NET1	SLC30A10
KCND3	CYP2A13
ALDOB	HCP5
EXOSC1	ZNF175
BTC	KHDC1L
NOD2	PTK7
LYZ	FSD1
CCDC28B	SEPP1
FAM198B	ANP32A /// ANP32D
ZMYND10	S1PR5
EHD2	TEP1
CST3	NVL
RAPGEF4	ABCB1 /// ABCB4
FCN1	GZMH
IL5RA	FBXO4
OCA2	CARD9
KYNU	MYBL1
C5AR1	TARP
ARHGAP32	HSPC157
FCGRT	RHCE /// RHD
COL6A1	LOC100287602
41892	GFI1
NOP16	FCGBP
SYK	MMP23A /// MMP23B
VPS13B	ZBTB6
APOBEC3F	XK
EVI5	B3GALT2
RERE	C14orf104

CLC	DLX5
FCGR2B	CYP7A1
MEIS3P1	SFRP5
CCDC144A	KLRG1
LILRA4	CRCT1
POSTN	LAG3
CHD7	AGPAT4
C6orf26 /// MSH5	RNF32
PTK2	TARP /// TRGV3 /// TRGV5

APPENDIX C. SUPPLEMENTARY INFORMATION CHAPTER 4

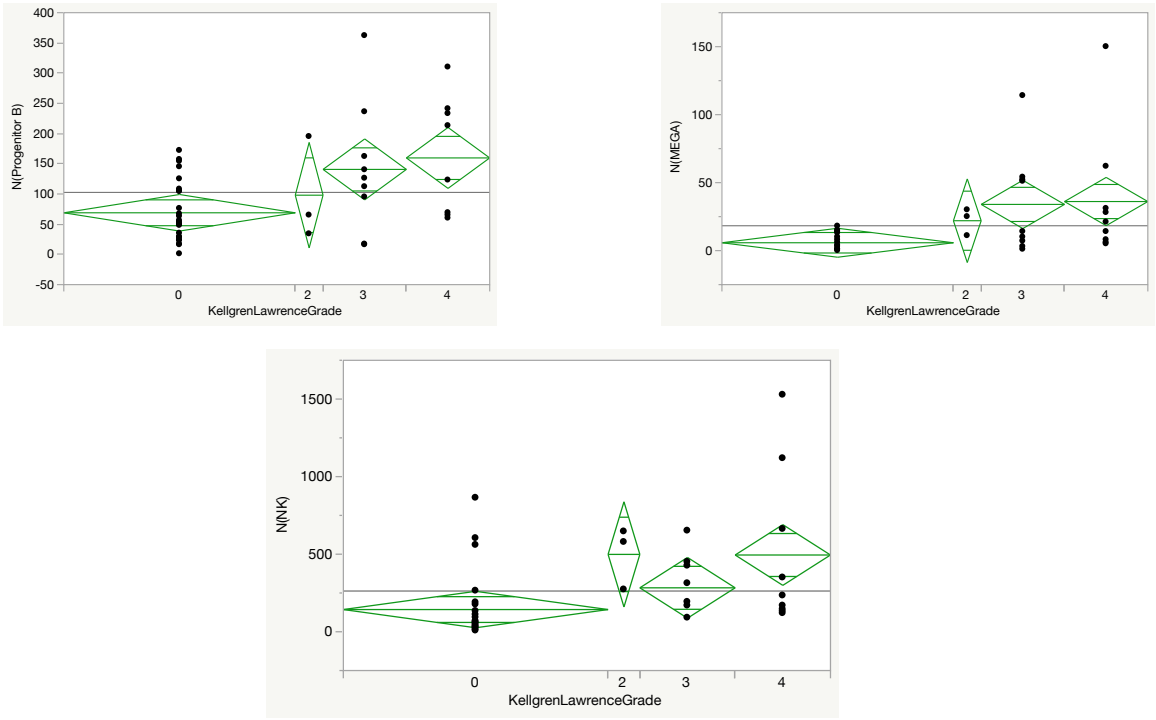


Figure S10. One-way anova on the cell ration of the cell types depending on the KL score. A significant change in the cell ratio was seen, depending on the KL score, for the progenitor B, MEGA and NK cells.

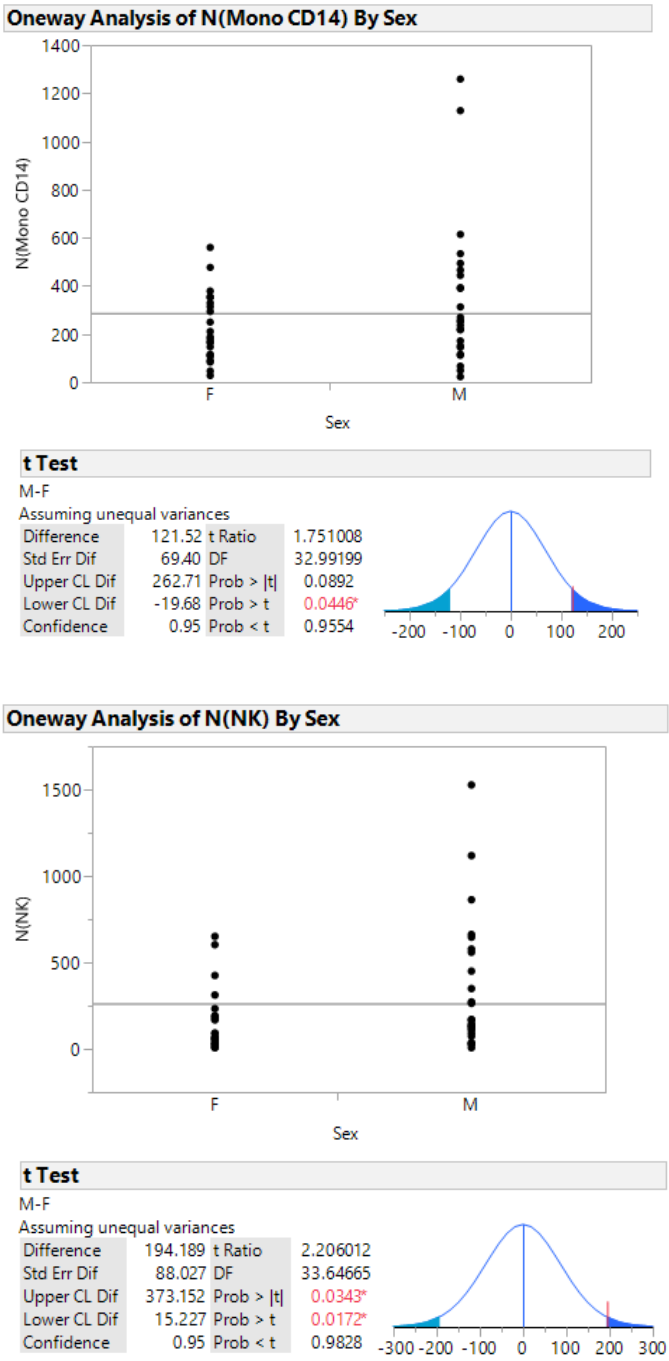


Figure S11. One-way anova on the cell ratio of the cell types depending on the sex. A significant difference was seen between male and female for the cell ratio of mono CD14 and NK cells.

PUBLICATIONS

1. Fernandes LM, Khan NM, **Trochez CM**, Duan M, Diaz-Hernandez ME, Presciutti SM, Gibson G, Drissi H. Single-cell RNA-seq identifies unique transcriptional landscapes of human nucleus pulposus and annulus fibrosus cells. Sci Rep. 2020 Sep 17;10(1):15263. doi: 10.1038/s41598-020-72261-7. PMID: 32943704; PMCID: PMC7499307.
2. Diaz-Hernandez ME, Khan NM, **Trochez CM**, Yoon T, Maye P, Presciutti SM, Gibson G, Drissi H. Derivation of notochordal cells from human embryonic stem cells reveals unique regulatory networks by single cell-transcriptomics. J Cell Physiol. 2020 Jun;235(6):5241-5255. doi: 10.1002/jcp.29411. Epub 2019 Dec 16. PMID: 31840817; PMCID: PMC7056550.
3. **Medrano-Trochez C**, Chatterjee P, Pradhan P, et al. Single-Cell RNAseq of Out-of-Thaw Mesenchymal Stromal Cells Shows Striking Tissue-of-Origin Differences and Inter-donor Cell-Cycle Variations. bioRxiv; 2020. DOI: 10.1101/2020.09.10.290155.
4. Pallab Pradhan, Paramita Chatterjee, Hazel Y. Stevens, Chad Glen, **Camila Medrano-Trochez**, Angela Jimenez, Linda Kippner, Wen Jun Seeto, Ye Li, Greg Gibson, Joanne Kurtzberg, Theresa Kontanchek, Carolyn Yeago, Krishnendu Roy. Single-Cell Transcriptomic Attributes and Unbiased Computational Modeling for the Prediction of Immunomodulatory Potency of Mesenchymal Stromal Cells. bioRxiv 2020.09.12.294850; doi: https://doi.org/10.1101/2020.09.12.294850

REFERENCES

REFERENCES

1. Larijani B, Esfahani EN, Amini P, Nikbin B, Alimoghaddam K, Amiri S, et al. Stem cell therapy in treatment of different diseases. Acta Med Iran. 2012;50(2):79-96.

2. Brown C, McKee C, Bakshi S, Walker K, Hakman E, Halassy S, et al. Mesenchymal stem cells: Cell therapy and regeneration potential. J Tissue Eng Regen Med. 2019;13(9):1738-55.

3. Galipeau J, Sensebe L. Mesenchymal Stromal Cells: Clinical Challenges and Therapeutic Opportunities. Cell Stem Cell. 2018;22(6):824-33.

4. Mastrolia I, Foppiani EM, Murgia A, Candini O, Samarelli AV, Grisendi G, et al. Challenges in Clinical Development of Mesenchymal Stromal/Stem Cells: Concise Review. Stem Cells Transl Med. 2019;8(11):1135-48.

5. Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, et al. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell. 2015;161(5):1202-14.

6. Zheng GX, Terry JM, Belgrader P, Ryvkin P, Bent ZW, Wilson R, et al. Massively parallel digital transcriptional profiling of single cells. Nat Commun. 2017;8:14049.

7. Phinney DG. Biochemical heterogeneity of mesenchymal stem cell populations: clues to their therapeutic efficacy. Cell Cycle. 2007;6(23):2884-9.

8. Diseases NNIoAaI. Overview of the Immune System 2013 [Available from: https://www.niaid.nih.gov/research/immune-system-overview.

9. Aristizábal B GÁ. Innate immune system. 2013

10. Blanchard N, Salvioni A, Robey EA. Chapter 26 - Adaptive immunity. In: Weiss LM, Kim K, editors. Toxoplasma gondii (Third Edition): Academic Press; 2020. p. 1107-46.

11. Janeway CA Jr TP, Walport M, et al. Immunobiology: The Immune System in Health and Disease. Science NYG, editor: The components of the immune system; 2001.

12. Autoimmune Diseases [Available from: https://www.niehs.nih.gov/health/topics/conditions/autoimmune/index.cfm#footnote1.

13. Ramos PS, Shedlock AM, Langefeld CD. Genetics of autoimmune diseases: insights from population genetics. J Hum Genet. 2015;60(11):657-64.

14. Meyer KC, Decker C, Baughman R. Toxicity and Monitoring of Immunosuppressive Therapy Used in Systemic Autoimmune Diseases. Clin Chest Med. 2010;31(3):565-+.

15. Ferrara JL, Levine JE, Reddy P, Holler E. Graft-versus-host disease. Lancet. 2009;373(9674):1550-61.

16. Billingham RE. The biology of graft-versus-host reactions. Harvey Lect. 1966;62:21-78.

17. Welniak LA, Blazar BR, Murphy WJ. Immunobiology of allogeneic hematopoietic stem cell transplantation. Annu Rev Immunol. 2007;25:139-70.

18. Loiseau P, Busson M, Balere ML, Dormoy A, Bignon JD, Gagne K, et al. HLA Association with hematopoietic stem cell transplantation outcome: the number of mismatches at HLA-A, -B, -C, -DRB1, or -DQB1 is strongly associated with overall survival. Biol Blood Marrow Transplant. 2007;13(8):965-74.

19. Shlomchik WD, Couzens MS, Tang CB, McNiff J, Robert ME, Liu J, et al. Prevention of graft versus host disease by inactivation of host antigen-presenting cells. Science. 1999;285(5426):412-5.

20. Reddy P, Maeda Y, Liu C, Krijanovski OI, Korngold R, Ferrara JL. A crucial role for antigen-presenting cells and alloantigen expression in graft-versus-leukemia responses. Nat Med. 2005;11(11):1244-9.

21. Dinse GE, Parks CG, Weinberg CR, Co CA, Wilkerson J, Zeldin DC, et al. Increasing Prevalence of Antinuclear Antibodies in the United States. Arthritis Rheumatol. 2020;72(6):1026-35.

22. Swift A. Osteoarthritis 1: Physiology, risk factors and causes of pain. Nurs Times. 2012;108(7):12-5.

23. Horvath G, Koroknai G, Acs B, Than P, Bellyei A, Illes T. Prevalence of radiographic primary hip and knee osteoarthritis in a representative Central European population. Int Orthop. 2011;35(7):971-5.

24. Salve H, Gupta V, Palanivel C, Yadav K, Singh B. Prevalence of knee osteoarthritis amongst perimenopausal women in an urban resettlement colony in South Delhi. Indian J Public Health. 2010;54(3):155-7.

25. Osteoarthritis (OA).

26. Dieppe P. Developments in osteoarthritis. Rheumatology. 2011;50(2):245-7.

27. Martel-Pelletier J, Pelletier JP. Is osteoarthritis a disease involving only cartilage or other articular tissues? Eklem Hast Cerrahisi. 2010;21(1):2-14.

28. Abramson SB, Attur M. Developments in the scientific understanding of osteoarthritis. Arthritis Research & Therapy. 2009;11(3).

29. Goldring MB, Goldring SR. Articular cartilage and subchondral bone in the pathogenesis of osteoarthritis. Ann Ny Acad Sci. 2010;1192:230-7.

30. Pollard TCB, Gwilym SE, Carr AJ. The assessment of early osteoarthritis. J Bone Joint Surg Br. 2008;90b(4):411-21.

31. Dequeker J, Luyten FP. The history of osteoarthritis-osteoarthrosis. Ann Rheum Dis. 2008;67(1):5-10.

32. Ding CH, Jones G, Wluka A, Cicuttini FM. What can we learn about osteoarthritis by studying a healthy person against a person with early onset of disease? Curr Opin Rheumatol. 2010;22(5):520-7.

33. Man GS, Mologhianu G. Osteoarthritis pathogenesis - a complex process that involves the entire joint. J Med Life. 2014;7(1):37-41.

34. Neogi T, Felson D, Niu J, Lynch J, Nevitt M, Guermazi A, et al. Cartilage loss occurs in the same subregions as subchondral bone attrition: a within-knee subregion-matched approach from the Multicenter Osteoarthritis Study. Arthritis Rheum. 2009;61(11):1539-44.

35. Sutton S, Clutterbuck A, Harris P, Gent T, Freeman S, Foster N, et al. The contribution of the synovium, synovial derived inflammatory cytokines and neuropeptides to the pathogenesis of osteoarthritis. Vet J. 2009;179(1):10-24.

36. Martel-Pelletier J, Pelletier JP. Is osteoarthritis a disease involving only cartilage or other articular tissues? Eklem Hastalik Cerrahisi. 2010;21(1):2-14.

37. Loeuille D, Chary-Valckenaere I, Champigneulle J, Rat AC, Toussaint F, Pinzano-Watrin A, et al. Macroscopic and microscopic features of synovial membrane inflammation in the osteoarthritic knee: correlating magnetic resonance imaging findings with disease severity. Arthritis Rheum. 2005;52(11):3492-501.

38. Hasegawa M, Segawa T, Maeda M, Yoshida T, Sudo A. Thrombin-cleaved osteopontin levels in synovial fluid correlate with disease severity of knee osteoarthritis. J Rheumatol. 2011;38(1):129-34.

39. Goldring MB, Goldring SR. Osteoarthritis. J Cell Physiol. 2007;213(3):626-34.

40. Heijink A, Gomoll AH, Madry H, Drobnic M, Filardo G, Espregueira-Mendes J, et al. Biomechanical considerations in the pathogenesis of osteoarthritis of the knee. Knee Surg Sports Traumatol Arthrosc. 2012;20(3):423-35.

41. Stannus O, Jones G, Cicuttini F, Parameswaran V, Quinn S, Burgess J, et al. Circulating levels of IL-6 and TNF-alpha are associated with knee radiographic osteoarthritis and knee cartilage loss in older adults. Osteoarthritis Cartilage. 2010;18(11):1441-7.

42. Englund M. The role of the meniscus in osteoarthritis genesis. Med Clin North Am. 2009;93(1):37-43, x.

43. Sun Y, Mauerhan DR, Kneisl JS, James Norton H, Zinchenko N, Ingram J, et al. Histological examination of collagen and proteoglycan changes in osteoarthritic menisci. Open Rheumatol J. 2012;6:24-32.

44. Grainger AJ, Rhodes LA, Keenan AM, Emery P, Conaghan PG. Quantifying peri-meniscal synovitis and its relationship to meniscal pathology in osteoarthritis of the knee. Eur Radiol. 2007;17(1):119-24.

45. Sun Y, Mauerhan DR, Honeycutt PR, Kneisl JS, Norton HJ, Zinchenko N, et al. Calcium deposition in osteoarthritic meniscus and meniscal cell culture. Arthritis Res Ther. 2010;12(2):R56.

46. Bhattacharyya T, Gale D, Dewire P, Totterman S, Gale ME, McLaughlin S, et al. The clinical importance of meniscal tears demonstrated by magnetic resonance imaging in osteoarthritis of the knee. J Bone Joint Surg Am. 2003;85(1):4-9.

47. Kim GB, Seo MS, Park WT, Lee GW. Bone Marrow Aspirate Concentrate: Its Uses in Osteoarthritis. Int J Mol Sci. 2020;21(9).

48. Gupta PK, Chullikana A, Rengasamy M, Shetty N, Pandey V, Agarwal V, et al. Efficacy and safety of adult human bone marrow-derived, cultured, pooled, allogeneic mesenchymal stromal cells (Stempeucel(R)): preclinical and clinical trial in osteoarthritis of the knee joint. Arthritis Res Ther. 2016;18(1):301.

49. Jo CH, Lee YG, Shin WH, Kim H, Chai JW, Jeong EC, et al. Intra-articular injection of mesenchymal stem cells for the treatment of osteoarthritis of the knee: a proof-of-concept clinical trial. Stem Cells. 2014;32(5):1254-66.

50. Yubo M, Yanyan L, Li L, Tao S, Bo L, Lin C. Clinical efficacy and safety of mesenchymal stem cell transplantation for osteoarthritis treatment: A meta-analysis. PLoS One. 2017;12(4):e0175449.

51. Chahla J, Alland JA, Verma NN. Bone Marrow Aspirate Concentrate for Orthopaedic Use. Orthop Nurs. 2018;37(6):379-81.

52. Chahla J, Mandelbaum BR. Biological Treatment for Osteoarthritis of the Knee: Moving from Bench to Bedside-Current Practical Concepts. Arthroscopy. 2018;34(5):1719-29.

53. Jones IA, Togashi R, Wilson ML, Heckmann N, Vangsness CT, Jr. Intra-articular treatment options for knee osteoarthritis. Nat Rev Rheumatol. 2019;15(2):77-90.

54. Chahla J, Dean CS, Moatshe G, Pascual-Garrido C, Serra Cruz R, LaPrade RF. Concentrated Bone Marrow Aspirate for the Treatment of Chondral Injuries and Osteoarthritis of the Knee: A Systematic Review of Outcomes. Orthop J Sports Med. 2016;4(1):2325967115625481.

55. Kuraitis DG, C. & Suuronen, E.J. & Ruel, M. Cell therapy to regenerate the ischemic heart. Cardiac Regeneration and Repair. 2014.

56. Facts about cell therapy 2020 [Available from: https://www.aabb.org/news-resources/resources/cellular-therapies/facts-about-cellular-therapies.

57. Orozco L, Munar A, Soler R, Alberca M, Soler F, Huguet M, et al. Treatment of knee osteoarthritis with autologous mesenchymal stem cells: a pilot study. Transplantation. 2013;95(12):1535-41.

58. Vangsness CT, Jr., Farr J, 2nd, Boyd J, Dellaero DT, Mills CR, LeRoux-Williams M. Adult human mesenchymal stem cells delivered via intra-articular injection to the knee following partial medial meniscectomy: a randomized, double-blind, controlled study. J Bone Joint Surg Am. 2014;96(2):90-8.

59. Kim C, Keating A. Cell Therapy for Knee Osteoarthritis: Mesenchymal Stromal Cells. Gerontology. 2019;65(3):294-8.

60. Lindvall O, Bjorklund A. Cell therapy in Parkinson's disease. NeuroRx. 2004;1(4):382-93.

61. Dominici M, Le Blanc K, Mueller I, Slaper-Cortenbach I, Marini F, Krause D, et al. Minimal criteria for defining multipotent mesenchymal stromal cells. The International Society for Cellular Therapy position statement. Cytotherapy. 2006;8(4):315-7.

62. Liu X, Xiang Q, Xu F, Huang J, Yu N, Zhang Q, et al. Single-cell RNA-seq of cultured human adipose-derived mesenchymal stem cells. Sci Data. 2019;6:190031.

63. Lo Surdo JL, Millis BA, Bauer SR. Automated microscopy as a quantitative method to measure differences in adipogenic differentiation in preparations of human mesenchymal stromal cells. Cytotherapy. 2013;15(12):1527-40.

64. Samsonraj RM, Rai B, Sathiyanathan P, Puan KJ, Rotzschke O, Hui JH, et al. Establishing criteria for human mesenchymal stem cell potency. Stem Cells. 2015;33(6):1878-91.

65. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet. 2009;10(1):57-63.

66. Evans TG. Considerations for the use of transcriptomics in identifying the 'genes that matter' for environmental adaptation. J Exp Biol. 2015;218(Pt 12):1925-35.

67. Loebel C, Burdick JA. Engineering Stem and Stromal Cell Therapies for Musculoskeletal Tissue Repair. Cell Stem Cell. 2018;22(3):325-39.

68. Kim H, Bae C, Kook YM, Koh WG, Lee K, Park MH. Mesenchymal stem cell 3D encapsulation technologies for biomimetic microenvironment in tissue regeneration. Stem Cell Res Ther. 2019;10(1):51.

69. Richardson SM, Kalamegam G, Pushparaj PN, Matta C, Memic A, Khademhosseini A, et al. Mesenchymal stem cells in regenerative medicine: Focus on articular cartilage and intervertebral disc regeneration. Methods. 2016;99:69-80.

70. Buenrostro JD, Corces MR, Lareau CA, Wu B, Schep AN, Aryee MJ, et al. Integrated Single-Cell Analysis Maps the Continuous Regulatory Landscape of Human Hematopoietic Differentiation. Cell. 2018;173(6):1535-48 e16.

71. Psaila B, Barkas N, Iskander D, Roy A, Anderson S, Ashley N, et al. Single-cell profiling of human megakaryocyte-erythroid progenitors identifies distinct megakaryocyte and erythroid differentiation pathways. Genome Biol. 2016;17:83.

72. Velten L, Haas SF, Raffel S, Blaszkiewicz S, Islam S, Hennig BP, et al. Human haematopoietic stem cell lineage commitment is a continuous process. Nat Cell Biol. 2017;19(4):271-81.

73. Papalexi E, Satija R. Single-cell RNA sequencing to explore immune cell heterogeneity. Nat Rev Immunol. 2018;18(1):35-45.

74. Villani AC, Satija R, Reynolds G, Sarkizova S, Shekhar K, Fletcher J, et al. Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. Science. 2017;356(6335).

75. Bjorklund AK, Forkel M, Picelli S, Konya V, Theorell J, Friberg D, et al. The heterogeneity of human CD127(+) innate lymphoid cells revealed by single-cell RNA sequencing. Nat Immunol. 2016;17(4):451-60.

76. Huang Y, Li Q, Zhang K, Hu M, Wang Y, Du L, et al. Single cell transcriptomic analysis of human mesenchymal stem cells reveals limited heterogeneity. Cell Death Dis. 2019;10(5):368.

77. Sun C, Wang L, Wang H, Huang T, Yao W, Li J, et al. Single-cell RNA-seq highlights heterogeneity in human primary Wharton's jelly mesenchymal stem/stromal cells cultured in vitro. Stem Cell Res Ther. 2020;11(1):149.

78. Barrett AN, Fong CY, Subramanian A, Liu W, Feng Y, Choolani M, et al. Human Wharton's Jelly Mesenchymal Stem Cells Show Unique Gene Expression Compared with Bone Marrow Mesenchymal Stem Cells Using Single-Cell RNA-Sequencing. Stem Cells Dev. 2019;28(3):196-211.

79. Khong SML, Lee M, Kosaric N, Khong DM, Dong Y, Hopfner U, et al. Single-Cell Transcriptomics of Human Mesenchymal Stem Cells Reveal Age-Related Cellular Subpopulation Depletion and Impaired Regenerative Function. Stem Cells. 2019;37(2):240-6.

80. Moll G, Geissler S, Catar R, Ignatowicz L, Hoogduijn MJ, Strunk D, et al. Cryopreserved or Fresh Mesenchymal Stromal Cells: Only a Matter of Taste or Key to Unleash the Full Clinical Potential of MSC Therapy? Adv Exp Med Biol. 2016;951:77-98.

81. Klein AM, Mazutis L, Akartuna I, Tallapragada N, Veres A, Li V, et al. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. Cell. 2015;161(5):1187-201.

82. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, 3rd, et al. Comprehensive Integration of Single-Cell Data. Cell. 2019;177(7):1888-902 e21.

83. Kiselev VY, Kirschner K, Schaub MT, Andrews T, Yiu A, Chandra T, et al. SC3: consensus clustering of single-cell RNA-seq data. Nat Methods. 2017;14(5):483-6.

84. Santos A, Wernersson R, Jensen LJ. Cyclebase 3.0: a multi-organism database on cell-cycle regulation and phenotypes. Nucleic Acids Res. 2015;43(Database issue):D1140-4.

85. Chinnadurai R, Copland IB, Garcia MA, Petersen CT, Lewis CN, Waller EK, et al. Cryopreserved Mesenchymal Stromal Cells Are Susceptible to T-Cell Mediated Apoptosis Which Is Partly Rescued by IFNgamma Licensing. Stem Cells. 2016;34(9):2429-42.

86. Regmi S, Pathak S, Kim JO, Yong CS, Jeong JH. Mesenchymal stem cell therapy for the treatment of inflammatory diseases: Challenges, opportunities, and future perspectives. Eur J Cell Biol. 2019;98(5-8):151041.

87. Ghannam S, Bouffi C, Djouad F, Jorgensen C, Noel D. Immunosuppression by mesenchymal stem cells: mechanisms and clinical applications. Stem Cell Res Ther. 2010;1(1):2.

88. Harris SG, Padilla J, Koumas L, Ray D, Phipps RP. Prostaglandins as modulators of immunity. Trends Immunol. 2002;23(3):144-50.

89. Aggarwal S, Pittenger MF. Human mesenchymal stem cells modulate allogeneic immune cell responses. Blood. 2005;105(4):1815-22.

90. Niehage C, Steenblock C, Pursche T, Bornhauser M, Corbeil D, Hoflack B. The cell surface proteome of human mesenchymal stromal cells. PLoS One. 2011;6(5):e20399.

91. Le Blanc K, Mougiakakos D. Multipotent mesenchymal stromal cells and the innate immune system. Nat Rev Immunol. 2012;12(5):383-96.

92. Oldenborg PA, Gresham HD, Lindberg FP. CD47-signal regulatory protein alpha (SIRPalpha) regulates Fcgamma and complement receptor-mediated phagocytosis. J Exp Med. 2001;193(7):855-62.

93. Jaiswal S, Jamieson CH, Pang WW, Park CY, Chao MP, Majeti R, et al. CD47 is upregulated on circulating hematopoietic stem cells and leukemia cells to avoid phagocytosis. Cell. 2009;138(2):271-85.

94. Toledano N, Gur-Wahnon D, Ben-Yehuda A, Rachmilewitz J. Novel CD47: SIRPalpha dependent mechanism for the activation of STAT3 in antigen-presenting cell. PLoS One. 2013;8(9):e75595.

95. Picarda E, Ohaegbulam KC, Zang X. Molecular Pathways: Targeting B7-H3 (CD276) for Human Cancer Immunotherapy. Clin Cancer Res. 2016;22(14):3425-31.

96. Vences-Catalan F, Rajapaksa R, Srivastava MK, Marabelle A, Kuo CC, Levy R, et al. Tetraspanin CD81 promotes tumor growth and metastasis by modulating the functions of T regulatory and myeloid-derived suppressor cells. Cancer Res. 2015;75(21):4517-26.

97. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26(1):139-40.

98. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102(43):15545-50.

99. Chen J, Bardes EE, Aronow BJ, Jegga AG. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. Nucleic Acids Res. 2009;37(Web Server issue):W305-11.

100. Huse M. Mechanical forces in the immune system. Nat Rev Immunol. 2017;17(11):679-90.

101. Simpson E. Special regulatory T-cell review: Regulation of immune responses--examining the role of T cells. Immunology. 2008;123(1):13-6.

102. Jamil MO, Mineishi S. State-of-the-art acute and chronic GVHD treatment. Int J Hematol. 2015;101(5):452-66.

103. U.S. Department of Health and Human Services, Office on Women's Health 2019 [Available from: https://www.womenshealth.gov/a-z-topics/autoimmune-diseases.

104. Sendi P, Wolf A, Graber P, Zimmerli W. Multiple opportunistic infections after high-dose steroid therapy for giant cell arteritis in a patient previously treated with a purine analog. Scand J Infect Dis. 2006;38(10):922-4.

105. Hu C, Li L. The immunoregulation of mesenchymal stem cells plays a critical role in improving the prognosis of liver transplantation. J Transl Med. 2019;17(1):412.

106. Burnham AJ, Daley-Bauer LP, Horwitz EM. Mesenchymal stromal cells in hematopoietic cell transplantation. Blood Adv. 2020;4(22):5877-87.

107. Zimmermann JA, Hettiaratchi MH, McDevitt TC. Enhanced Immunosuppression of T Cells by Sustained Presentation of Bioactive Interferon-gamma Within Three-Dimensional Mesenchymal Stem Cell Constructs. Stem Cells Transl Med. 2017;6(1):223-37.

108. English K, Barry FP, Field-Corbett CP, Mahon BP. IFN-gamma and TNF-alpha differentially regulate immunomodulation by murine mesenchymal stem cells. Immunol Lett. 2007;110(2):91-100.

109. Kim DS, Jang IK, Lee MW, Ko YJ, Lee DH, Lee JW, et al. Enhanced Immunosuppressive Properties of Human Mesenchymal Stem Cells Primed by Interferon-gamma. EBioMedicine. 2018;28:261-73.

110. Liang C, Jiang E, Yao J, Wang M, Chen S, Zhou Z, et al. Interferon-gamma mediates the immunosuppression of bone marrow mesenchymal stem cells on T-lymphocytes in vitro. Hematology. 2018;23(1):44-9.

111. Lin T, Pajarinen J, Nabeshima A, Lu L, Nathan K, Jamsen E, et al. Preconditioning of murine mesenchymal stem cells synergistically enhanced immunomodulation and osteogenesis. Stem Cell Res Ther. 2017;8(1):277.

112. Guess AJ, Daneault B, Wang R, Bradbury H, La Perle KMD, Fitch J, et al. Safety Profile of Good Manufacturing Practice Manufactured Interferon gamma-Primed Mesenchymal Stem/Stromal Cells for Clinical Trials. Stem Cells Transl Med. 2017;6(10):1868-79.

113. Wang M, Yuan Q, Xie L. Mesenchymal Stem Cell-Based Immunomodulation: Properties and Clinical Application. Stem Cells Int. 2018;2018:3057624.

114. Kovach TK, Dighe AS, Lobo PI, Cui Q. Interactions between MSCs and immune cells: implications for bone healing. J Immunol Res. 2015;2015:752510.

115. Le Blanc K, Davies LC. Mesenchymal stromal cells and the innate immune response. Immunol Lett. 2015;168(2):140-6.

116. Aune TM, Pierce CW. Activation of a suppressor T-cell pathway by interferon. Proc Natl Acad Sci U S A. 1982;79(12):3808-12.

117. Jung MK, Lee JS, Kwak JE, Shin EC. Tumor Necrosis Factor and Regulatory T Cells. Yonsei Med J. 2019;60(2):126-31.

118. Ding Y, Bushell A, Wood KJ. Mesenchymal stem-cell immunosuppressive capabilities: therapeutic implications in islet transplantation. Transplantation. 2010;89(3):270-3.

119. Keating A. Mesenchymal stromal cells. Curr Opin Hematol. 2006;13(6):419-25.

120. Chang JW, Tsai HL, Chen CW, Yang HW, Yang AH, Yang LY, et al. Conditioned mesenchymal stem cells attenuate progression of chronic kidney disease through inhibition of epithelial-to-mesenchymal transition and immune modulation. J Cell Mol Med. 2012;16(12):2935-49.

121. Bohacova P, Holan V. Mesenchymal stem cells and type 1 diabetes treatment. Vnitr Lek.64(7-8):725-8.

122. Klinker MW, Marklein RA, Lo Surdo JL, Wei CH, Bauer SR. Morphological features of IFN-gamma-stimulated mesenchymal stromal cells predict overall immunosuppressive capacity. Proc Natl Acad Sci U S A. 2017;114(13):E2598-E607.

123. Vellasamy S, Sandrasaigaran P, Vidyadaran S, Abdullah M, George E, Ramasamy R. Mesenchymal stem cells of human placenta and umbilical cord suppress T-cell proliferation at G0 phase of cell cycle. Cell Biol Int. 2013;37(3):250-6.

124. Lee S, Kim S, Chung H, Moon JH, Kang SJ, Park CG. Mesenchymal stem cell-derived exosomes suppress proliferation of T cells by inducing cell cycle arrest through p27kip1/Cdk2 signaling. Immunol Lett. 2020;225:16-22.

125. Li X, Xu Z, Bai J, Yang S, Zhao S, Zhang Y, et al. Umbilical Cord Tissue-Derived Mesenchymal Stem Cells Induce T Lymphocyte Apoptosis and Cell Cycle Arrest by Expression of Indoleamine 2, 3-Dioxygenase. Stem Cells Int. 2016;2016:7495135.

126. Taechangam N, Iyer SS, Walker NJ, Arzi B, Borjesson DL. Mechanisms utilized by feline adipose-derived mesenchymal stem cells to inhibit T lymphocyte proliferation. Stem Cell Res Ther. 2019;10(1):188.

127. Andrews S. FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. 2010.

128. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1):15-21.

129. Shen J, Abu-Amer Y, O'Keefe RJ, McAlinden A. Inflammation and epigenetic regulation in osteoarthritis. Connect Tissue Res. 2017;58(1):49-63.

130. Griffin TM, Scanzello CR. Innate inflammation and synovial macrophages in osteoarthritis pathophysiology. Clin Exp Rheumatol. 2019;37 Suppl 120(5):57-63.

131. Scanzello CR. Role of low-grade inflammation in osteoarthritis. Curr Opin Rheumatol. 2017;29(1):79-85.

132. Marchev AS, Dimitrova PA, Burns AJ, Kostov RV, Dinkova-Kostova AT, Georgiev MI. Oxidative stress and chronic inflammation in osteoarthritis: can NRF2 counteract these partners in crime? Ann N Y Acad Sci. 2017;1401(1):114-35.

133. Centeno C, Pitts J, Al-Sayegh H, Freeman M. Efficacy of autologous bone marrow concentrate for knee osteoarthritis with and without adipose graft. Biomed Res Int. 2014;2014:370621.

134. Kim JD, Lee GW, Jung GH, Kim CK, Kim T, Park JH, et al. Clinical outcome of autologous bone marrow aspirates concentrate (BMAC) injection in degenerative arthritis of the knee. Eur J Orthop Surg Traumatol. 2014;24(8):1505-11.

135. Mautner K, Bowers R, Easley K, Fausel Z, Robinson R. Functional Outcomes Following Microfragmented Adipose Tissue Versus Bone Marrow Aspirate Concentrate Injections for Symptomatic Knee Osteoarthritis. Stem Cells Transl Med. 2019;8(11):1149-56.

136. Pittenger MF, Mackay AM, Beck SC, Jaiswal RK, Douglas R, Mosca JD, et al. Multilineage potential of adult human mesenchymal stem cells. Science. 1999;284(5411):143-7.

137. Bain BJ. The bone marrow aspirate of healthy subjects. Br J Haematol. 1996;94(1):206-9.

138. Kumagai H, Yoshioka T, Sugaya H, Tomaru Y, Shimizu Y, Yamazaki M, et al. Quantitative assessment of mesenchymal stem cells contained in concentrated autologous bone marrow aspirate transplantation for the treatment of osteonecrosis of the femoral head: predictive factors and differences by etiology. BMC Res Notes. 2018;11(1):848.

139. Baxter MA, Wynn RF, Jowitt SN, Wraith JE, Fairbairn LJ, Bellantuono I. Study of telomere length reveals rapid aging of human marrow stromal cells following in vitro expansion. Stem Cells. 2004;22(5):675-82.

140. Centeno C, Sheinkop M, Dodson E, Stemper I, Williams C, Hyzy M, et al. A specific protocol of autologous bone marrow concentrate and platelet products versus exercise therapy for symptomatic knee osteoarthritis: a randomized controlled trial with 2 year follow-up. J Transl Med. 2018;16(1):355.

141. Centeno CJ, Al-Sayegh H, Bashir J, Goodyear S, Freeman MD. A dose response analysis of a specific bone marrow concentrate treatment protocol for knee osteoarthritis. BMC Musculoskelet Disord. 2015;16:258.

142. Oetjen KA, Lindblad KE, Goswami M, Gui G, Dagur PK, Lai C, et al. Human bone marrow assessment by single-cell RNA sequencing, mass cytometry, and flow cytometry. JCI Insight. 2018;3(23).

143. Diekman BO, Guilak F. Stem cell-based therapies for osteoarthritis: challenges and opportunities. Curr Opin Rheumatol. 2013;25(1):119-26.

144. Troeberg L, Nagase H. Proteases involved in cartilage matrix degradation in osteoarthritis. Biochim Biophys Acta. 2012;1824(1):133-45.

145. Yasuda T. Cartilage destruction by matrix degradation products. Mod Rheumatol. 2006;16(4):197-205.

146. Narayan A, Berger, B, Cho, H. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. 2018.

147. NIH. Autoimmune diseases. 2020.

148. NIH. Cance types 2020 [Available from: https://www.cancer.gov/types.

149. Clinic M. Osteoarthritis 2020 [Available from: https://www.mayoclinic.org/diseases-conditions/osteoarthritis/diagnosis-treatment/drc-20351930.

150. NIH. Types of cancer treatment. 2020.

151. Dazzi F, van Laar JM, Cope A, Tyndall A. Cell therapy for autoimmune diseases. Arthritis Res Ther. 2007;9(2):206.

152. Rad F, Ghorbani M, Mohammadi Roushandeh A, Habibi Roudkenar M. Mesenchymal stem cell-based therapy for autoimmune diseases: emerging roles of extracellular vesicles. Mol Biol Rep. 2019;46(1):1533-49.

153. Swart JF, Delemarre EM, van Wijk F, Boelens JJ, Kuball J, van Laar JM, et al. Haematopoietic stem cell transplantation for autoimmune diseases. Nat Rev Rheumatol. 2017;13(4):244-56.

154. El-Badawy A, El-Badri N. Clinical Efficacy of Stem Cell Therapy for Diabetes Mellitus: A Meta-Analysis. PLoS One. 2016;11(4):e0151938.

155. Ismail A, Sharrack B, Saccardi R, Moore JJ, Snowden JA. Autologous haematopoietic stem cell therapy for multiple sclerosis: a review for supportive care clinicians on behalf of the Autoimmune Diseases Working Party of the European Society

for Blood and Marrow Transplantation. Curr Opin Support Palliat Care. 2019;13(4):394-401.

156. Buzhor E, Leshansky L, Blumenthal J, Barash H, Warshawsky D, Mazor Y, et al. Cell-based therapy approaches: the hope for incurable diseases. Regen Med. 2014;9(5):649-72.

157. Alexander T, Farge D, Badoglio M, Lindsay JO, Muraro PA, Snowden JA, et al. Hematopoietic stem cell therapy for autoimmune diseases - Clinical experience and mechanisms. J Autoimmun. 2018;92:35-46.